Goodness of fit for lattice processes

Hidalgo, Javier

Accepted Manuscript

Goodness of fit for lattice processes

Javier Hidalgo

PII: S0304-4076(09)00076-1
DOI: 10.1016/j.jeconom.2009.03.003
Reference: ECONOM 3186

To appear in: *Journal of Econometrics*

Received date: 10 November 2008
Accepted date: 13 March 2009

Please cite this article as: Hidalgo, J., Goodness of fit for lattice processes. *Journal of Econometrics* (2009), doi:10.1016/j.jeconom.2009.03.003

GOODNESS OF FIT FOR LATTICE PROCESSES

JAVIER HIDALGO

Abstract. The paper discusses tests for the correct specification of a model when data is observed in a d-dimensional lattice, extending previous work when the data is collected in the real line. As it happens with the latter type of data, the asymptotic distribution of the tests are functionals of a Gaussian sheet process, say $\boldsymbol{B}(\nu)$, $\nu \in [0,\pi]^d$. Because it is not easy to find a time transformation $h(\nu)$ such that $\boldsymbol{B}(h(\nu))$ becomes the standard Brownian sheet, a consequence is that the critical values are difficult, if at all possible, to obtain. So, to overcome the problem of its implementation, we propose to employ a bootstrap approach, showing its validity in our context.

JEL Classification: C21, C23.

1. INTRODUCTION

The paper is concerned with testing the goodness of fit of a parametric family of models for data collected in a lattice. More specifically, we are concerned with the correct specification (or model selection) of the dynamic structure with time series and/or spatial stationary processes $\{x(t)\}_{t\in\mathbb{Z}}$ defined on a d-dimensional lattice. The key idea of the test is to compare how close is the parametric and nonparametric fits of the data to provide support for the null hypothesis. In the paper, we shall explicitly consider data for which $d \leq 3$. The motivation lies in the fact that the most often type of data available in economics is when $d = 2$, say with agricultural or environmental data, or when $d = 3$. An important example of the latter is the spatial-temporal data sets, that is data collected in a lattice during a number of periods. However, we ought to mention that extensions to higher index lattice processes can be adapted under suitable modifications.

All throughout the paper we will assume that the (spatial) process $\{x(t)\}_{t\in\mathbb{Z}^d}$ can be represented by the multilateral model

$$x(t) - \mu = \sum_{j\in\mathbb{Z}^d} \psi(j)\,\varepsilon(t-j), \quad \sum_{j\in\mathbb{Z}^d} \psi^2(j) < \infty \quad \psi(0) = 1, \tag{1.1}$$

for some sequence $\{\varepsilon(t)\}_{t\in\mathbb{Z}^d}$ satisfying $\mathbb{E}(\varepsilon(t)) = 0$ and $\mathbb{E}(\varepsilon(0)\,\varepsilon(t)) = \sigma_\varepsilon^2 \mathcal{I}(t=0)$, where $\mathcal{I}(\cdot)$ denotes the indicator function. Notice that because our model is multilateral, the sequence $\{\varepsilon(t)\}_{t\in\mathbb{Z}^d}$ loses its interpretation as the "prediction" error or that it can be regarded as innovations. Under (1.1), the spectral density function of $\{x(t)\}_{t\in\mathbb{Z}^d}$, $f(\lambda)$, can be factorized as

$$f(\lambda) = \frac{\sigma_\varepsilon^2}{(2\pi)^d} |\Psi(\lambda)|^2, \quad \lambda \in \Pi^d,$$

where $\Pi = (-\pi, \pi]$ and with $\Psi(\lambda) = \sum_{j\in\mathbb{Z}^d} \psi(j)\, e^{-ij\cdot\lambda}$. The function $\Psi(\lambda)$ summarizes the covariogram structure of $\{x(t)\}_{t\in\mathbb{Z}^d}$, which is the main feature to obtain

Date: 10 November 2008.

Key words and phrases. Goodness of fit tests. Spatial linear processes. Spectral domain. Bootstrap tests.

The paper has benefited from the comments of two referees. Of course, any remaining errors are the author's sole responsibility.

accurate prediction/extrapolation and/or interpolation (kriging) in the case of spatial data. Notice that the ultimate aim when modelling data is nothing but to predict the future. Henceforth the notation "$j \cdot \lambda$" means the inner product of multi-indices j and λ of dimensional d. Also, any element a that belongs to $\mathbb{Z}^d$ (or Π^d), the d-fold Cartesian product of the set $\mathbb{Z}$ (or Π), will be referred to as a multi-index of dimension d. Also, we shall write, say, $a = (a[1], ..., a[d])$ with the square brackets used to denote the components of a.

The aim of the paper is to test whether the data support the null hypothesis that $\Psi(\lambda)$ belongs to a specific parametric family

$$\mathcal{H} = \{\Psi_\theta(\lambda) : \theta \in \Theta\}, \tag{1.2}$$

where $\Theta \subset \mathbb{R}^p$ is a proper compact parameter set. That is, we are interested on the null hypothesis

$$H_0 : \forall\lambda \in [-\pi, \pi]^d \text{ and for some } \theta_0 \in \Theta, \;\; |\Psi(\lambda)|^2 = |\Psi_{\theta_0}(\lambda)|^2. \tag{1.3}$$

The alternative hypothesis H_1 is the negation of H_0. Alternatively we could have formulated the null hypothesis in terms of the covariogram $\{\gamma(s)\}_{s\in\mathbb{Z}^d}$, where $\gamma(s) = Cov(x(t), x(t+s))$. That is, the null hypothesis is that the covariogram follows a particular parametric family, say $\{\gamma(s)\}_{s\in\mathbb{Z}^d} = \{\gamma_\vartheta(s)\}_{s\in\mathbb{Z}^d}$, where from now on we denote $\vartheta = (\theta', \sigma_\varepsilon^2)'$. This is the case after observing that for any stationary spatial lattice process $\{x(t)\}_{t\in\mathbb{Z}^d}$, the spectral density $f(\lambda)$ and the covariogram $\gamma(s)$ are related through the expression

$$\begin{matrix} \gamma(s) \\ \gamma_\vartheta(s) \end{matrix} = \left\{ \begin{matrix} \int_{\Pi^d} f(\lambda) e^{-is\cdot\lambda} d\lambda \\ \int_{\Pi^d} f_\vartheta(\lambda) e^{-is\cdot\lambda} d\lambda \end{matrix} \right. ; \quad s \in \mathbb{Z}^d.$$

One parameterization of (1.1), or (2.12) below, is the $ARMA$ field model

$$P(L)(x(t) - \mu) = Q(L)\varepsilon(t),$$

where denoting henceforth for muti-indeces z and j, $z^j = \prod_{\ell=1}^d z[\ell]^{j[\ell]}$ with the convention that $0^0 = 1$,

$$P(z) = \sum_{j\in\mathbb{Z}^d} \alpha(j) z^j; \quad \alpha(0) = 1; \quad Q(z) = \sum_{j\in\mathbb{Z}^d} \beta(j) z^j; \quad \beta(0) = 1,$$

are finite series in $\mathbb{Z}^d$. That is, only a finite number of the $\alpha(j)'s$ and $\beta(j)'s$ coefficients are non-zero. For instance the $ARMA$ field model given by

$$\sum_{j=-k_1}^{k_2} \alpha(j)(x(t-j) - \mu) = \sum_{j=-\ell_1}^{\ell_2} \beta(j)\varepsilon(t-j) \qquad \alpha(0) = \beta(0) = 1$$

whose spectral density function is

$$f(\lambda) = \frac{\sigma_\varepsilon^2}{(2\pi)^d} \frac{\left|\sum_{j=-\ell_1}^{\ell_2} \beta(j) e^{ij\cdot\lambda}\right|^2}{\left|\sum_{j=-k_1}^{k_2} \alpha(j) e^{ij\cdot\lambda}\right|^2}.$$

Notice that the $ARMA$ field model becomes a causal representation if the polynomials $Q(L)$ and $P(L)$ are both unilateral. It is worth mentioning that Whittle (1954) showed that, almost any given stationary bilateral scheme on a plane lattice, there corresponds a unilateral autoregression having the same spectral scheme although not necessarily of finite order as is the case when $d = 1$. See also Guyon (1982a).

Another parametric model of interest is the extension of the classical Bloomfield (1973) exponential model, see also Whittle's (1954) Section 6, to processes in a lattice. These models are characterized as having a spectral density function

$$f_{\vartheta}(\lambda) = \sigma_{\varepsilon}^{2} \exp\left\{ -\sum_{\ell \prec 0} \alpha(\ell;\theta) \cos(\ell \cdot \lambda) \right\},$$

where "$\prec$" denotes the lexicographical (dictionary) ordering which is defined as

$$j \prec k \Leftrightarrow (\exists \ell > 0)(\forall i < \ell)(j[i] = k[i] \wedge j[\ell] < k[\ell]),$$

that is, if one of the terms $j[\ell] < k[\ell]$ and all the proceeding ones are equal. For instance, when $d = 2$, we would then have that, say, $\ell \prec 0$ corresponds to the half plane of $\mathbb{Z}^2$, $\overline{\mathbb{Z}}^2 = \left\{ (\ell[1], \ell[2]) \in \mathbb{Z}^d : (\ell[1] = 0 \wedge \ell[2] < 0) \vee \ell[1] < 0 \right\}$. Observe that if we allowed ℓ in the last displayed equality to belong to $\mathbb{Z}^d$ the model would not be then identified as $\cos(\ell \cdot \lambda) = \cos(-\ell \cdot \lambda)$.

When $d = 1$, the problem of testing a specific dependence structure of the data is very exhaustive and prominent. Different tests have been formulated using either the spectral density or the autocorrelation functions. Regarding the former, we can cite among others, the pioneer work by Grenander and Rosenblatt (1957) to test for the null hypothesis of white noise dependence. A classical test using the autocorrelation function is the Box and Pierce (1970) statistic. For a latter reference, see Delgado, Hidalgo and Velasco (2005) and references therein. In the paper, we have chosen to employ frequency domain techniques or to base the test in terms of the spectral density function, contrary to a "time domain" approach based on the covariance/variogram structure of the data.

Our tests fall into the category of goodness of fit tests as we do not specify any particular alternative model or family. The tests are based on a direct comparison between two estimates of the spectral density function in a way similar to the well known Hausman-Durbin-Wu's test. That is, they rely on the comparison of two estimates: one which is only consistent under the null, whereas the second (less efficient) estimator is consistent under the maintained hypothesis. Although the literature when $d > 1$ is not very vast and exhaustive, some work has already been done, see for instance Diblasi and Bowman (2001) or Crujeiras et al. (2006). However, our work differs from theirs in that contrary to Diblasi and Bowman (2001) we do test for general specifications and that contrary to Crujeiras et al. (2006) our test does not involve any bandwidth or smoothing parameter. In fact, the latter approach uses the distance between a smooth estimator of the spectral density function and its parametric estimator under H_0. This approach provides asymptotically distribution free tests under suitable conditions on the smoothing parameter, see for instance Hong (1996) or Paparoditis (2000) among others. However, the latter approach seems to be a mere artifact when testing for a particular parametric family and the final outcome of all these tests may depend on the arbitrary choice of the bandwidth parameter(s) for which no relevant theory is available for testing purposes. That is, there are not rules available on how to choose the bandwidth parameter with empirical data. In fact, we might face the strange situation that with the same data set two different practitioners might conclude differently. The latter is clearly not very desirable from both theoretical or applied stand point of view. So, in this context, one of our main motivation is to extend goodness-of-fit tests examined and described when $d = 1$ to $d \geq 1$, where we do not require the choice of any bandwidth parameter. For that purpose, we rely on the periodogram which although it is not a consistent estimator for $f(\lambda)$, its integral is a consistent estimator of the spectral distribution function as the integral is the most natural smoothing algorithm.

The remainder of the paper is organized as follows. In the next section, we present the test and examine its asymptotic properties when the true value of the parameter θ_0 is known, whereas Section 3 extends these results to more realistic situations where we need to estimate the parameters of the model. Because, the asymptotic distribution of the test in the latter scenario is not pivotal and model dependent, Section 4 describes the bootstrap test showing its validity. Section 5 gives the proof of a series of lemmas employed in the proof of our main results in Section 6.

2. TESTS WHEN THE PARAMETERS ARE KNOWN

This section discusses and examines how we can test the null hypothesis H_0 given in (1.3). That is, the hypothesis

$$H_0 : f(\lambda) = \frac{\sigma_\varepsilon^2}{(2\pi)^d} \left|\Psi_{\theta_0}(\lambda)\right|^2 \quad \forall \lambda \in \widetilde{\Pi}^d \text{ for some value } \theta_0,$$

when the "true" value of θ_0 is known, and where herewith $\widetilde{\Pi}^d$ denotes $[0,\pi] \times [-\pi,\pi]^{d-1}$, that is $\lambda \in \widetilde{\Pi}^d$ if $\lambda[1] \in [0,\pi]$ and $\lambda[\ell] \in [-\pi,\pi]$ for $\ell = 2, ..., d$. Before we introduce and describe the test, we notice that we can alternatively state the null hypothesis H_0 as

$$H_0 : \frac{G_{\theta_0}(\lambda)}{G_{\theta_0}(\pi)} = \prod_{\ell=1}^{d} \frac{\lambda[\ell]}{\pi} \quad \text{for all } \lambda \in [0,\pi]^d, \tag{2.1}$$

where $G_\theta(\lambda) = 2\int_{-\lambda}^{\lambda} |\Psi_\theta(\omega)|^{-2} f(\omega)\, d\omega$ with the notation

$$\int_{\mu}^{\lambda} = \int_{(\mu[1]\wedge 0)}^{\lambda[1]} \int_{\mu[2]}^{\lambda[2]} \cdots \int_{\mu[d]}^{\lambda[d]}. \tag{2.2}$$

Under H_0, $G_{\theta_0}(\lambda)$ is the spectral distribution function of the lattice process $\{\varepsilon(t)\}_{t\in\mathbb{Z}}$ and $G_{\theta_0}(\pi) = \sigma_\varepsilon^2$. Notice that by symmetry of $f(\lambda)$, it is irrelevant which coordinate we choose to belong only to $[0,\pi]$ as the choice does not affect the value of $G_\theta(\lambda)$ and so the value of the test given below.

Given a record $\{x(t)\}_{t=1}^{n}$ and denoting henceforth $N = \prod_{\ell=1}^{d} n[\ell]$, a natural estimator of $G_{\theta_0}(\lambda)$ is

$$\widetilde{G}_{\theta,N}(\lambda) = 2\frac{1}{N} \sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]} \frac{I_x(\lambda_j)}{|\Psi_\theta(\lambda_j)|^2},$$

where, for a generic sequence $\{v(t)\}_{t=1}^{n}$, $I_v(\lambda)$ denotes the *periodogram*

$$I_v(\lambda) = \frac{1}{N}\left|\sum_{t=1}^{n} v(t)\, e^{-it\cdot\lambda}\right|^2; \qquad \lambda \in \widetilde{\Pi}^d$$

and similarly to the definition of $\int_\mu^\lambda$, we employ henceforth the notation

$$\sum_{j=[\tilde{n}\nu/\pi]}^{[\tilde{n}\mu/\pi]} = \sum_{j[1]=[\tilde{n}[1]\nu[1]/\pi]_+}^{[\tilde{n}[1]\mu[1]/\pi]} \sum_{j[2]=[\tilde{n}[2]\nu[2]/\pi]}^{[\tilde{n}[2]\mu[2]/\pi]} \cdots \sum_{j[d]=[\tilde{n}[d]\nu[d]/\pi]}^{[\tilde{n}[d]\mu[d]/\pi]}, \tag{2.3}$$

where $[q]_+ = \max\{|q|, 1\}$. Also we have abbreviated $[n[\ell]/2]$ by $\tilde{n}[\ell]$ for $\ell = 1, ..., d$. As usual we have excluded the frequency $\lambda_j = 0$ from the sum $\sum_{j=[\tilde{n}\nu/\pi]}^{[\tilde{n}\mu/\pi]}$, so that we can take $Ex(t) = 0$ or assume that $x(t)$ has been already centered around its sample mean. It is often the case that in real applications, in order to make use of the *fast Fourier transform*, the periodogram is evaluated at the Fourier frequencies, that is

$\lambda_k = \left(\lambda_{k[1]}, ..., \lambda_{k[d]}\right)'$, where with $k[1] = 0, 1, ..., \tilde{n}[1]$ and $k[\ell] = 0, \pm 1, ..., \pm \tilde{n}[\ell]$ for $\ell = 2, ..., d$,

$$\lambda_{k[\ell]} = \frac{2\pi k[\ell]}{n[\ell]}; \quad \ell = 1, ..., d.$$

Unfortunately, as noted by Guyon (1982*b*), due to nonnegligible end effects, the bias of $I_x(\lambda_j)$ does not converge to zero fast enough when $d > 1$, so that it would have unwanted consequences. One of these is that the Whittle estimator of ϑ, see Guyon (1982*b*), does not have the standard asymptotic properties as when $d = 1$. Because of that, in the paper, we shall employ the taper periodogram defined as

$$I_v^T(\lambda_j) = \left|w_v^T(\lambda_j)\right|^2, \tag{2.4}$$

for a generic sequence $\{v(t)\}_{t=1}^n$ and

$$w_v^T(\lambda_j) = \frac{1}{\left(\sum_{t=1}^n h^2(t)\right)^{1/2}} \sum_{t=1}^n h(t)\, v(t)\, e^{it\cdot\lambda_j}$$

is the taper discrete Fourier transform. Tapering is primarily a technique employed to reduce the bias of the "standard" periodogram $I_v(\lambda)$. Notice that when $h(t) = 1$, we have that the *taper discrete Fourier transform* $w_v^T(\lambda_j)$ becomes the standard discrete Fourier transform *(DFT)*. It is worth mentioning that to alleviate the bias problem, alternative procedures to tapering have been proposed. One of these proposals was due to Guyon (1982*b*), who replaced the periodogram by

$$I_v^*(\lambda_k) = \frac{1}{(2\pi)^d} \sum_{h\in\mathcal{D}} \widehat{\gamma}_v^*(h)\, e^{-ih\cdot\lambda_k},$$

where $\widehat{\gamma}_v^*(h) = \frac{1}{N-|h|}\sum_{t(h)} v(t)\, v(t+h)$ and $\mathcal{D} = \{h : -n[\ell] < h[\ell] < n[\ell];\ \ell = 1, ..., d\}$. However, Dahlhaus and Künsch (1987) have criticized the use of $I_v^*(\lambda_k)$ on the grounds that the Whittle estimator, see (3.1) below, loses its minimum distance interpretability and that the objective function possesses several local maxima. The latter implies that to obtain the maximum of the Whittle function becomes more strenuous. Another possibility is the one described by Robinson and Vidal-Sanz (2006). The latter proposal will be helpful when $d \geq 4$. However as we only consider explicitly the most common scenario $d \leq 3$, it suffices for our results to hold true to employ the taper periodogram $I_v^T(\lambda_j)$.

The benefits of tapering can be seen following the properties of the *cosine-bell (or Hanning)* taper, which is defined as

$$h(t) = \frac{1}{2^d}\prod_{\ell=1}^d h_\ell(t[\ell]); \qquad h_\ell(t[\ell]) = \left(1 - \cos\left(\frac{2\pi t[\ell]}{n[\ell]}\right)\right). \tag{2.5}$$

Indeed, denoting the taper Dirichlet kernel by $D_\ell^T(\lambda[\ell]) = \sum_{t[\ell]=1}^{n[\ell]} h_\ell(t[\ell])\, e^{it[\ell]\lambda[\ell]}$, we have that

$$\sup_{n[\ell],\lambda[\ell]>0}\left|D_\ell^T(\lambda[\ell])\right| = O\left(\min\left\{n[\ell], n[\ell]^{-2}|\lambda[\ell]|^{-3}\right\}\right). \tag{2.6}$$

The immediate consequence of property (2.6) is that the bias of the taper periodogram is of smaller order of magnitude than the one of the standard periodogram. Observe that

$$D_\ell^T\left(\lambda_{j[\ell]}\right) = \frac{1}{6^{1/2}}\left\{-D_\ell\left(\lambda_{j[\ell]-1}\right) + 2D_\ell\left(\lambda_{j[\ell]}\right) - D_\ell\left(\lambda_{j[\ell]+1}\right)\right\}, \tag{2.7}$$

where $D_\ell(\lambda[\ell]) = \sum_{t[\ell]=1}^{n[\ell]} e^{it[\ell]\lambda[\ell]}$ is the Dirichlet kernel. It is worth observing that the standard *DFT* and the *cosine-bell* taper *DFT* are related by the equality

$$w_x^T(\lambda_j) = \frac{1}{6^{d/2}} \prod_{\ell=1}^{d} \left[-w_x\left(\lambda_{j[\ell]-1}\right) + 2w_x\left(\lambda_{j[\ell]}\right) - w_x\left(\lambda_{j[\ell]+1}\right) \right]. \tag{2.8}$$

In the paper we shall explicitly consider the *cosine-bell*, although the same results follow employing other taper functions such as *Parzen or Kolmogorov's* tapers.

The formulation of H_0 given in (2.1) suggests to use the Bartlett's $T_p - process$ as a basis for testing H_0. The $T_p - process$ is defined as

$$\alpha_{\theta,N}(\lambda) = 2^{-1/2} N^{1/2} \left[\frac{G_{\theta,N}(\lambda)}{G_{\theta,N}(\pi)} - \prod_{\ell=1}^{d} \left(\frac{\lambda[\ell]}{\pi} \right) \right], \quad \lambda \in [0,\pi]^d, \tag{2.9}$$

where

$$G_{\theta,N}(\lambda) = 2\frac{1}{N} \sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]} \frac{I_x^T(\lambda_j)}{|\Psi_\theta(\lambda_j)|^2}. \tag{2.10}$$

It is worth mentioning that similarly we might have employed the $U_p - process$ as Grenander and Rosenblatt (1957) did. The latter is defined as

$$U_{\theta,N}(\lambda) = 2\frac{1}{N} \sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]} \left\{ I_x^T(\lambda_j) - \sigma_\varepsilon^2 |\Psi_\theta(\lambda_j)|^2 \right\}.$$

One motivation to employ $\alpha_{\theta,N}(\lambda)$ instead of $U_{\theta,N}(\lambda)$ is that the latter statistic is not invariant to the variance of $\{\varepsilon(t)\}_{t\in\mathbb{Z}^d}$ as is the former statistic $\alpha_{\theta,N}(\lambda)$ in (2.9). Notice that because we have excluded the frequency $\lambda_j = 0$ from the definition of $\sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]}$ and $\alpha_{\theta,N}(\lambda)$ is scale invariant, it is easy to show that a linear transformation of the data does not change the value of $\alpha_{\theta,N}$ and therefore we can assume, without loss of generality, that $\mathbb{E}x(t) = 0$ and $\mathrm{Var}(\varepsilon(t)) = 1$.

One rational of the statistic $\alpha_{\theta,N}(\lambda)$ follows from the observation (see Lemma 4 in Section 5) that under H_0, we have that

$$\max_{-\tilde{n}\leq j\leq \tilde{n}} \mathbb{E}\left| \frac{I_x^T(\lambda_j)}{|\Psi_{\theta_0}(\lambda_j)|^2} - I_\varepsilon^T(\lambda_j) \right| = o(1),$$

where "$a \leq b$" means that $a[\ell] \leq b[\ell]$ for all $\ell = 1,...,d$. Also, observe that $0 < j[1] \leq \tilde{n}[1]$ whereas $-\tilde{n}[\ell] < j[\ell] \leq \tilde{n}[\ell]$ for $\ell = 2,...,d$.

Thus, from the previous observation, we can expect that $\alpha_{\theta_0,N}$ will be asymptotically equivalent to Bartlett's $U_p - process$ for $\{\varepsilon(t)\}_{t\in\mathbb{Z}^d}$, i.e.

$$\alpha_N^0(\lambda) = 2^{-1/2} N^{1/2} \left[\frac{G_N^0(\lambda)}{G_N^0(\pi)} - \prod_{\ell=1}^{d} \left(\frac{\lambda[\ell]}{\pi} \right) \right], \tag{2.11}$$

with $G_N^0(\lambda) = 2\frac{1}{N}\sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]} I_\varepsilon^T(\lambda_j)$, $\lambda \in [0,\pi]^d$. Observe that the $U_p - process$ α_N^0 and the $T_p - process$ $\alpha_{\theta_0,N}$ are identical when $\{x(t)\}_{t\in\mathbb{Z}^d}$ is a "white noise" lattice process.

Let us introduce the following regularity conditions.

Condition C1: (a) $\{\varepsilon(t)\}_{t\in\mathbb{Z}^d}$ in (1.1) is a zero mean independent identically distributed sequence of random variables with variance $\sigma_\varepsilon^2 = 1$ and finite $4th$ moments with κ_ε denoting the fourth cumulant of $\{\varepsilon(t)\}_{t\in\mathbb{Z}^d}$.

(b) The multilateral *Moving Average* representation of $\{x(t)\}_{t\in\mathbb{Z}^d}$ in (1.1) can be written (or it has a representation) as a multilateral *Autocorrelation* model

$$\sum_{j\in\mathbb{Z}^d} \xi(j)\, x(t-j) = \varepsilon(t) \qquad \xi(0)=1, \tag{2.12}$$

where $\xi(j)$ is the coefficient of z^j in the Fourier expansion of $\mathcal{L}^{-1}(z)$, where

$$\mathcal{L}(z) = \mathcal{L}(z[1],...,z[d]) = \sum_{j\in\mathbb{Z}^d} \psi(j)\, z^j.$$

Condition C2: $N = \prod_{\ell=1}^{d} n[\ell]$, where $n[\ell] \asymp \vec{n}$ for $\ell = 1,...,d$, and "$a \asymp b$" means that $C^{-1} \leq a/b \leq C$ for some finite positive constant C.

Condition C3: $\{h(t)\}_{t=1}^{n}$ is the cosine-bell taper function in (2.5).

We now comment on Conditions $C1$ to $C3$. Part (a) of Condition $C1$ seems to be a minimal condition for Proposition 1 below to hold true. Observe that due to the quadratic nature of α_N^0, for the latter to have finite second moments, we require finite fourth moments of $\{\varepsilon(t)\}_{t\in\mathbb{Z}^d}$. Also we have assumed that the true value of σ_ε^2 is 1. The latter follows from our comments made after the definition of $G_{\theta,N}(\lambda)$ in (2.10). However, we shall emphasize that we are not saying or suggesting that the true value of σ_ε^2 is known, only that it is equal to 1. Sufficient regularity conditions required for the validity of the expansion in (2.12) is $\Psi(z)$ be no zero for any $z[\ell]$, $\ell = 1,...d$, which simultaneously satisfy $|z[1]| = 1,...,|z[d]| = 1$ at least when the *Moving Average* representation is of finite order. The latter implies that $f(\lambda)$ is a positive function.

Looking at the proof of Proposition 1 below, and then that of Theorem 1, it appears that we do not need to assume finite fourth moments of the sequence $\{\varepsilon(t)\}_{t\in\mathbb{Z}}$. The reason is similar to the work of Anderson and Walker (1964). However, as in the more realistic situation when we need to estimate the unknown parameters of the model, we require finite fourth moments to obtain the asymptotic properties of the estimates, we have just preferred to leave the condition as it stands.

Condition $C2$ can be generalized to the case where the rate of convergence to zero of $n^{-1}[\ell]$ differs for different $\ell = 1,...,d$. However, for notational simplicity we prefer to leave it as it stands. On the other hand, in $C3$ the taper function employed for the asymptotics to follow can be more general, as those given by Kolmogorov's or Parzen's tapers. In fact, in situations where $d > 3$, it might be needed for the results of the paper to follow. However, as the most important cases in empirical applications are covered in the paper, we shall leave the cosine-bell taper explicitly as the taper function to be employed.

The empirical processes $\alpha_N^0(\lambda)$ and $\alpha_{\theta_0,N}(\lambda)$ given in (2.11) and (2.9) respectively are random elements in $D[0,\pi]^d$. The functional space $D[0,\pi]^d$ is endowed with the Skorohod's metric (see Billingsley, 1968 or Bickel and Wichura, 1971) and convergence in distribution in the corresponding topology is denoted by "$\Rightarrow$".

Proposition 1. *Under $C1 - C3$, we have that*

$$\alpha_N^0(\lambda) \Rightarrow \widetilde{\boldsymbol{B}}(\lambda) = \boldsymbol{B}\left(\frac{\lambda}{\pi}\right) - \prod_{\ell=1}^{d}\left(\frac{\lambda[\ell]}{\pi}\right)\boldsymbol{B}(1) \qquad \lambda \in [0,\pi]^d, \tag{2.13}$$

where $\left\{\boldsymbol{B}(u) : u \in [0,1]^d\right\}$ is the standard Brownian sheet.

Remark 1. *Recall that the covariance structure of the standard Brownian sheet is*

$$Cov(\boldsymbol{B}(u), \boldsymbol{B}(v)) = \prod_{\ell=1}^{d}(u[\ell] \wedge v[\ell]), \quad \textit{for } u,v \in [0,1]^d.$$

Proposition 1 extends Grenander and Rosenblatt's (1957) results when $d = 1$, although under stronger conditions than the ones we have assumed in this paper. In particular, we do not need to assume eighth bounded moments.

To establish the asymptotic equivalence between $\alpha_{\theta_0,N}$ and α_N^0, we introduce the following smoothness condition.

Condition C.4: $|\Psi(\lambda)|^2 = \left|\sum_{j\in\mathbb{Z}^d} \psi(j)\, e^{-ij\cdot\lambda}\right|^2$ is a positive and continuously differentiable function on $[-\pi,\pi]^d$.

Our main result of this section is the following theorem.

Theorem 1. *Consider* (1.1) *and assume* $C1 - C4$*. Then, under* H_0*,*

$$\alpha_{\theta_0,N}(\lambda) \Rightarrow \widetilde{\boldsymbol{B}}(\lambda) \qquad \lambda \in [0,\pi]^d .$$

Proof. The proof is an immediate consequence of Proposition 1 and Lemma 4 after we observe that Lemma 4, with $\zeta(\lambda) = 1$ there, implies that

$$N^{1/2} \sup_{\lambda\in[0,\pi]^d} \left|G_{\theta_0,N}(\lambda) - G_N^0(\lambda)\right| = o_p(1)$$

so that $N^{1/2} \sup_{\lambda\in[0,\pi]^d} \left|\frac{G_{\theta_0,N}(\lambda)}{G_{\theta_0,N}(\pi)} - \frac{G_N^0(\lambda)}{G_N^0(\pi)}\right| = o_p(1)$ by standard algebra. □

Remark 2. *An immediate conclusion from Theorem 1 and Proposition 1 is that*

$$G_{\theta_0,N}(\pi) - \sigma_\varepsilon^2 = O_p\left(N^{-1/2}\right). \tag{2.14}$$

We now comment on the result of Theorem 1. The theorem indicates that $\alpha_{\theta_0,N}$ is asymptotically pivotal. One consequence is that critical regions of tests based on a continuous functional $\eta : D[0,\pi]^d \mapsto \mathbb{R}^+$ can be easily obtained. Different functionals η lead to tests with different power properties. Among them are omnibus, directional and/or Portmanteau-type tests. For example, classical functionals which lead to omnibus tests are the Kolmogorov-Smirnov ($\eta(g) = \sup_{\lambda\in[0,\pi]^d} |g(\lambda)|$) and the Cramér-von Mises ($\eta(g) = \pi^{-d} \int_{-\pi}^{\pi} g(\lambda)^2\, d\lambda$).

In fact we have the following corollary.

Corollary 1. *Under* H_0 *and* $C1 - C4$*, we have that for any continuous functional* $\eta(\cdot)$*,* $\eta(\alpha_{\theta_0,N}) \xrightarrow{d} \eta\left(\widetilde{\boldsymbol{B}}\right)$*.*

Proof. The proof follows from Theorem 1 and the continuous mapping theorem. □

Unfortunately, the results of Theorem 1 and Corollary 1 are only valid when the "true" value of θ_0 is known, which in practical situations is unrealistic. The question is then how are our previous results affected when we estimate θ_0? This is the topic of the next section.

3. TESTS WHEN THE PARAMETERS ARE UNKNOWN

This section extends the results of Section 2 to the more realistic situation where we need to estimate the parameters θ_0 to implement the test. That is, we replace θ_0 in $\alpha_{\theta,N}(\lambda)$ by an estimator, for example $\widehat{\theta}$ given in (3.1) below. In this scenario, drawing the terminology from Durbin (1973), we say that our null hypothesis H_0 becomes a composite hypothesis.

A popular estimator of $\vartheta_0' = \left(\theta_0', \sigma_\varepsilon^2\right)$ is the Whittle (1954) estimator defined as

$$\widehat{\vartheta}_c = \arg \min_{\vartheta\in\Theta\times\mathbb{R}^+} \mathcal{Q}^c(\vartheta),$$

where $\mathcal{Q}^c(\vartheta) = \int_{-\pi}^{\pi} \left\{ \log f_\vartheta(\lambda) + \frac{I_x^T(\lambda)}{(2\pi)^d f_\vartheta(\lambda)} \right\} d\lambda$ or in its discrete version

$$\widehat{\vartheta} = \arg \min_{\vartheta \in \Theta \times \mathbb{R}^+} \mathcal{Q}_N(\vartheta), \tag{3.1}$$

where

$$\mathcal{Q}_N(\vartheta) = \frac{1}{N} \sum_{j=-\tilde{n}}^{\tilde{n}} \left\{ \log f_\vartheta(\lambda_j) + \frac{I_x^T(\lambda_j)}{(2\pi)^d f_\vartheta(\lambda_j)} \right\} \tag{3.2}$$

with $f_\vartheta(\lambda_j) = \sigma_\varepsilon^2 |\Psi_\theta(\lambda_j)| / (2\pi)^d$ and $\Theta \subset \mathbb{R}^p$ is a compact set. Recall our notation given in (2.3) and that the true value of the variance of $\varepsilon(t)$ is unknown.

In this scenario, the $T_p - process$ $\alpha_{\theta_0,N}(\lambda)$ becomes

$$\alpha_{\widehat{\theta},N}(\lambda) = 2^{-1/2} N^{1/2} \left[\frac{G_{\widehat{\theta},N}(\lambda)}{G_{\widehat{\theta},N}(\pi)} - \prod_{\ell=1}^{d} \left(\frac{\lambda[\ell]}{\pi} \right) \right], \ \ \lambda \in [0,\pi]^d, \tag{3.3}$$

where $G_{\theta,N}(\lambda)$ is given in (2.10).

It is worth noticing that, contrary to the standard causal models, as Whittle (1954) first noticed, the estimator of ϑ_0 obtained by

$$\mathring{\theta} = \arg\min_{\theta \in \Theta} \frac{2}{N} \sum_{j=-\tilde{n}}^{\tilde{n}} \frac{I_x^T(\lambda_j)}{|\Psi_\theta(\lambda_j)|^2}, \quad \mathring{\sigma}_\varepsilon^2 = \frac{2}{N} \sum_{j=-\tilde{n}}^{\tilde{n}} \frac{I_x^T(\lambda_j)}{|\Psi_{\mathring{\theta}}(\lambda_j)|^2}$$

is inconsistent. The main reason for the lack of consistency of $\mathring{\theta}$ is that when the model is not causal $\int_{-\pi}^{\pi} \varphi_\theta(\lambda) d\lambda \neq 0$, where from now on we write

$$\varphi_\theta(\lambda) = \frac{\partial}{\partial \theta} \log |\Psi_\theta(\lambda)|^2 \tag{3.4}$$

and $\phi_\vartheta(\lambda) = \frac{\partial}{\partial \vartheta} \log f_\vartheta(\lambda) = \left(\varphi'_\theta(\lambda), \sigma_\varepsilon^{-2} \right)'$.

Let's introduce the following regularity conditions.

Condition C5: θ_0 is an interior point of the compact parameter set $\Theta \subset \mathbb{R}^p$.

Condition C6: $|\Psi_\theta(\lambda)|$ is a positive and twice continuously differentiable function in θ on $[-\pi,\pi]^d$, and continuously differentiable function on $[-\pi,\pi]^d$ for all $\theta \in \Theta$.

Condition C7: If $\theta_1 \neq \theta_2$, then $\Psi_{\theta_1}(\lambda) \neq \Psi_{\theta_2}(\lambda)$ in a set $\Delta \subset [-\pi,\pi]^d$ with positive Lebesgue measure.

The conditions imposed on Θ and the model (1.1) or (2.12) are standard so that we omit any comment on them. Let

$$q_{\vartheta,N} = \frac{1}{N} \sum_{j=-\tilde{n}}^{\tilde{n}} \phi_\vartheta(\lambda_j) \left\{ \frac{I_x^T(\lambda_j)}{\sigma_\varepsilon^2 |\Psi_\theta(\lambda_j)|^2} - 1 \right\}; \quad Q_{\vartheta,N} = \frac{1}{N} \sum_{j=-\tilde{n}}^{\tilde{n}} \phi_\vartheta(\lambda_j) \phi'_\vartheta(\lambda_j), \tag{3.5}$$

and also, recalling our notation in (2.2),

$$\Phi_\vartheta = (2\pi)^{-d} \int_{-\pi}^{\pi} \phi_\vartheta(\lambda) d\lambda \ \ \text{and} \ \ \Lambda_\vartheta = (2\pi)^{-d} \int_{-\pi}^{\pi} \phi_\vartheta(\lambda) \phi'_\vartheta(\lambda) d\lambda.$$

Notice that we write explicitly σ_ε^2 as it is a parameter in itself.

Condition C8: Λ_{ϑ_0} is a continuously positive definite matrix.

Theorem 2. *Under C1-C3 and $C5 - C8$, we have that*

$$N^{1/2} \left(\widehat{\vartheta} - \vartheta_0 \right) \xrightarrow{d} \mathcal{N}\left(0, 2\Lambda_{\vartheta_0}^{-1} V_{\vartheta_0} \Lambda_{\vartheta_0}^{-1} \right),$$

where $V_{\vartheta_0} = 2\Lambda_{\vartheta_0} + \kappa_\varepsilon (35/18)^d \Phi_{\vartheta_0} \Phi'_{\vartheta_0}$.

Proof. First, by definition, we know that

$$\widehat{\vartheta}-\vartheta_0=-\overline{Q}_{\widetilde{\vartheta},N}^{-1}q_{\vartheta_0,N},$$

where $\widetilde{\vartheta}$ is an intermediate point between ϑ_0 and $\widehat{\vartheta}$, $q_{\vartheta,N}$ is given in (3.5) and $\overline{Q}_{\vartheta,N}$ is given by

$$\begin{aligned} & Q_{\vartheta,N}+\frac{1}{N}\sum_{j=-\tilde{n}}^{\tilde{n}}\left(2\phi_{\vartheta}\left(\lambda_j\right)\phi_{\vartheta}'\left(\lambda_j\right)-\frac{\partial^2 f_{\vartheta}\left(\lambda_j\right)}{\partial\vartheta\partial\vartheta'}\right)\left\{\frac{I_x^T\left(\lambda_j\right)}{\sigma_{\varepsilon}^2\left|\Psi_{\theta}\left(\lambda_j\right)\right|^2}-1\right\} \\ = \;& Q_{\vartheta,N}+o_p\left(1\right) \end{aligned}$$

by Lemma 5 and that $\widehat{\vartheta}-\vartheta_0=o_p\left(1\right)$ by Lemma 6. On the other hand, by Brillinger (1981, p.15) and standard arguments, since $\widetilde{\vartheta}-\vartheta_0=o_p\left(1\right)$, we have that $Q_{\widetilde{\vartheta},N}-\Lambda_{\vartheta_0}=o_p\left(1\right)$. Next, by Lemma 4 with $\zeta\left(\lambda\right)=\phi_{\vartheta_0}\left(\lambda_j\right)$ there,

$$q_{\vartheta_0,N}=\frac{1}{N}\sum_{j=-\tilde{n}}^{\tilde{n}}\phi_{\vartheta_0}\left(\lambda_j\right)\left\{I_{\varepsilon}^T\left(\lambda_j\right)-1\right\}+o_p\left(1\right).$$

From here the proof proceeds as in Robinson and Vidal-Sanz (2006). □

Looking at the proof of Theorem 2, and denoting in what follows

$$\begin{aligned} \widetilde{\varphi}_{\theta}\left(\lambda\right) &= \varphi_{\theta}\left(\lambda\right)-\frac{2}{\left(2\pi\right)^d}\int_{-\pi}^{\pi}\varphi_{\theta}\left(\lambda\right)d\lambda, \quad \widetilde{\phi}_{\vartheta}\left(\lambda\right)=\left(\widetilde{\varphi}_{\theta}'\left(\lambda\right),0\right)' \\ \widetilde{\varphi}_{\theta,N}\left(\lambda_j\right) &= \varphi_{\theta}\left(\lambda_j\right)-\frac{2}{N}\sum_{j=-\tilde{n}}^{\tilde{n}}\varphi_{\theta}\left(\lambda_j\right), \quad \widetilde{\phi}_{\vartheta,N}\left(\lambda\right)=\left(\widetilde{\varphi}_{\theta,N}'\left(\lambda\right),0\right)' \end{aligned}$$

with $\varphi_{\theta}\left(\lambda\right)$ given in (3.4), standard algebra establishes that the Whittle estimator $\widehat{\vartheta}$ in (3.1) satisfies the asymptotic linearization

$$\widehat{\vartheta}-\vartheta_0=-Q_{\vartheta_0,N}^{-1}\left\{\int_{-\pi}^{\pi}\widetilde{\phi}_{\theta_0}\left(\lambda\right)\alpha_{\theta_0,N}\left(d\lambda\right)+\int_{-\pi}^{\pi}\phi_{\vartheta_0}\left(\lambda\right)d\lambda\frac{1}{N}\sum_{j=-\tilde{n}}^{\tilde{n}}\left(\frac{I_x^T\left(\lambda_j\right)}{\left(2\pi\right)^d f_{\vartheta_0}\left(\lambda_j\right)}-1\right)\right\}$$

(3.6)

$$+\,o_p\left(N^{-1/2}\right).$$

Now using (3.6) and defining

$$\alpha_{\infty}\left(\lambda\right)=\widetilde{\boldsymbol{B}}\left(\lambda\right)-\left(\frac{1}{\left(2\pi\right)^d}\int_{-\lambda}^{\lambda}\widetilde{\varphi}_{\theta_0}'\left(\bar{\lambda}\right)d\bar{\lambda}\right)\widetilde{\Lambda}^{-1}\left(\theta_0\right)\int_{-\pi}^{\pi}\widetilde{\varphi}_{\theta_0,N}\left(\bar{\lambda}\right)\widetilde{\boldsymbol{B}}\left(d\bar{\lambda}\right),$$

where $\widetilde{\Lambda}_{\theta}=\frac{1}{(2\pi)^d}\int_{-\pi}^{\pi}\widetilde{\varphi}_{\theta}\left(\bar{\lambda}\right)\widetilde{\varphi}_{\theta}'\left(\bar{\lambda}\right)d\bar{\lambda}$, we obtain the following result.

Theorem 3. *Under H_0 and assuming C1 − C3 and C5 − C8, uniformly in $\lambda\in[0,\pi]$,*

$$\begin{aligned} (a)\;\; \alpha_{\widehat{\theta},N}\left(\lambda\right) &= \alpha_N^0\left(\lambda\right)-\left(\frac{1}{N}\sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]}\widetilde{\varphi}_{\theta_0,N}'\left(\lambda_j\right)\right)\widetilde{\Lambda}_{\theta_0,N}^{-1}\frac{1}{N}\sum_{j=-\tilde{n}}^{\tilde{n}}\widetilde{\varphi}_{\theta_0,N}'\left(\lambda_j\right)I_{\varepsilon}^T\left(\lambda_j\right) \\ &\quad +o_p\left(1\right), \end{aligned}$$

and $\widetilde{\Lambda}_{\theta,N}=\frac{1}{N}\sum_{j=-\tilde{n}}^{\tilde{n}}\widetilde{\varphi}_{\theta,N}\left(\lambda_j\right)\widetilde{\varphi}_{\theta,N}'\left(\lambda_j\right)$.

(*b*) $\alpha_{\widehat{\theta},N}\Rightarrow\alpha_{\infty}$.

Corollary 2. *Let* $\widehat{\eta}_N := \eta\left(\alpha_{\widehat{\theta},N}\right)$, *where* $\eta(\cdot)$ *be a continuous functional* $\eta : D[0,\pi]^d \to \mathbb{R}$. *Under* H_0 *and the same conditions of Theorem 3, we have that*

$$\widehat{\eta}_N \xrightarrow{d} \eta(\alpha_\infty).$$

Proof. The proof follows from Theorem 3 and the continuous mapping theorem. □

The main conclusion that we draw from Theorem 3 is that the T_p-process $\alpha_{\widehat{\theta},N}$ is no longer asymptotically pivotal, so that the immediate consequence is that tests based on η, for example the Kolmogorov or Cramér-von Mises's statistics, are not useful for practical purposes as its asymptotic critical values are difficult, if at all possible, to obtain. To compute the critical values of the asymptotic distribution of $\widehat{\eta}_N$, several approaches have been described and examined. A first approach makes use of a bandwidth parameter that must behave in some required sense. This procedure makes the asymptotic distribution of the statistics $\widehat{\eta}_N$ pivotal, so that its critical values are readily available. Among them, the popular Portmanteau test. Box and Pierce (1970) showed that the partial sum of the residuals squared autocorrelations of a stationary ARMA process is approximately chi-squared distributed assuming that the number of autocorrelations considered diverges to infinity with the sample size at an appropriate rate. Alternatively we could employ a frequency domain approach as in Hong (1996) or Paparoditis (2000), who compared a nonparametric estimator of $f(\lambda)$ with the parametric one. The first shortcoming of the latter method is that the power of the test is smaller than the one proposed in the paper, that is if we denote by b_N the bandwidth parameter, their test has a local power of order $\left(Nb_N^d\right)^{-1/2}$ whereas ours is $N^{-1/2}$. A second potential drawback is that the choice of b_N seems an artifact when testing for a particular parametric family and the final outcome of all these tests may depend on the arbitrary choice of the bandwidth parameter for which no relevant theory is available. That is, there are not rules available on how to choose b_N for the purpose of testing.

A second alternative is in the spirit of Durbin, Knott and Taylor (1976) for the classical empirical process, and it was the route followed by Anderson (1997), who proposed to approximate the critical value of the Cramér-von Mises test for a stationary AR model. The method considers a truncated version of the spectral representation of $\alpha_{\widehat{\theta},N}$ with estimated orthogonal components. The number of estimated orthogonal components must suitably increase with the sample size. However, its implementation is quite cumbersome even for the rather simpler case when $d = 1$. See for instance Anderson (1997) for details.

So, in view of the preceding arguments, we consider a third approach based on bootstrap algorithms. This is the route employed, among others, by Chen and Romano (2000) or Hainz and Dahlhaus (2000) for short-range models using the $U_p - process$ and by Hidalgo and Kreiss (2006), who allow also long-range dependence using the $T_p - process$. Of course all those articles were for $d = 1$. Also, we will see that bootstraps employed when $d = 1$ are not valid in our context.

4. BOOTSTRAP TEST FOR THE TEST

Since Efron (1979), bootstrap algorithms have become a common tool in applied work and thus considerable effort has been devoted to its development. The primary motivation for this effort is that they have proved to be a very useful statistical tool. We can cite two main examples/reasons. First, bootstrap methods are capable of approximating the finite sample distribution of statistics better than those based on their asymptotic counterparts. And secondly, and perhaps the most important, they allow computing valid asymptotic quantiles of the limiting distribution in situations when the practitioner is unable to compute its quantiles.

In the present paper we face the latter situation. Following our comments at the end of the previous section, the aim of this section is to propose a bootstrap procedure for $\alpha_{\widehat{\theta},N}$ given in (3.3) and thus for $\widehat{\eta}_N = \eta\left(\alpha_{\widehat{\theta},N}\right)$. The resampling method must be such that the bootstrap statistic, say $\widehat{\eta}_N^*$, is such that $\widehat{\eta}_N^* \to_{d^*} \eta(\alpha_\infty)$ in probability under H_0, where "$\to_{d^*}$" denotes

$$\Pr\left[\widehat{\eta}_N^* \leq z | \underset{\sim}{x}\right] \xrightarrow{p} \mathcal{G}(z),$$

at each continuity point z of $\mathcal{G}(z) = \Pr(\eta(\alpha_\infty) \leq z)$. Moreover, under local alternatives

$$H_a: \quad f_\vartheta(\lambda)\left(1 + \frac{1}{N^{1/2}} g(\lambda)\right) \text{ for some } \vartheta \in \Theta \times \mathbb{R}^+ \tag{4.1}$$

where $g(\lambda)$ is some symmetric, non-constant continuous function in $[0,\pi]^d$ such that $\frac{1}{N^{1/2}} g(\lambda) > -1$ for all $N \geq 1$, $\widehat{\eta}_N^*$ must also converge, in bootstrap distribution to $\eta(\alpha_\infty)$, whereas under the alternative hypothesis, we only require that $\widehat{\eta}_N^*$ is bounded in probability to have good power properties.

Remark 3. *We should point out that H_a could have been written as*

$$H_a: \quad f_\vartheta(\lambda) + \frac{1}{N^{1/2}} \widetilde{g}(\lambda) \text{ for some } \vartheta \in \Theta \times \mathbb{R}^+$$

where $\widetilde{g}(\lambda)$ is a positive integrable function. However, since we are concerned with the relative error of $I_x^T(\lambda_j)$ compared to $f_\vartheta(\lambda_j)\left(|\Psi_\theta(\lambda_j)|^2\right)$, we found notationally more convenient to write H_a as given in (4.1).

When $d = 1$, Hidalgo and Kreiss (2006) examined a bootstrap algorithm based on an approach in Hidalgo (2003) showing its validity and consistency. This bootstrap consists on the following *3 STEPS*.

STEP 1: Let $\widetilde{x}(t) = (x(t) - \overline{x}) / \widehat{\sigma}_x$, where $\overline{x} = N^{-1} \sum_{t=1}^n x(t)$ and $\widehat{\sigma}_x^2 = N^{-1} \sum_{t=1}^N (x(t) - \overline{x})^2$, and a random sample of size N with replacement from the empirical distribution of $\widetilde{x}(t)$, denoted by $\underset{\sim}{x}^* = \{x^*(t)\}_{t=1}^n$.

STEP 2: For $j = 1, ..., \tilde{n}$, compute the bootstrap periodogram

$$\widetilde{I}_{x^*}^T(\lambda_j) = f_{\widehat{\vartheta}}(\lambda_j)\, \check{I}_{x^*}^T(\lambda_j),$$

where $f_{\widehat{\vartheta}}(\lambda_j) = \frac{G_{\widehat{\theta},N}(\pi)}{(2\pi)^d} \left|\Psi_{\widehat{\theta}}(\lambda_j)\right|^2$ and $\check{I}_{x^*}^T(\lambda_j)$ as defined in (2.4) and then the bootstrap analogue of $\widehat{\vartheta}$ by

$$\check{\vartheta}^* = \arg \min_{\vartheta \in \Theta \times \mathbb{R}^+} \widetilde{\mathcal{Q}}_N^*(\vartheta), \tag{4.2}$$

where, with $f_\vartheta(\lambda_j) = \sigma_\varepsilon^2 |\Psi_\theta(\lambda_j)| / (2\pi)^d$,

$$\widetilde{\mathcal{Q}}_N^*(\vartheta) = \frac{1}{N} \sum_{j=-\tilde{n}}^{\tilde{n}} \left\{\log f_\vartheta(\lambda_j) + \frac{\widetilde{I}_{x^*}^T(\lambda_j)}{(2\pi)^d f_\vartheta(\lambda_j)}\right\}.$$

STEP 3: Compute the bootstrap $T_p - process$

$$\alpha_{\check{\theta}^*,N}(\lambda) = 2^{-1/2} N^{1/2} \left[\frac{\widetilde{G}_{\check{\theta}^*,N}^*(\lambda)}{G_{\check{\theta}^*,N}(\pi)} - \prod_{\ell=1}^d \left(\frac{\lambda[\ell]}{\pi}\right)\right], \ \ \lambda \in [0,\pi]^d,$$

where $\widetilde{G}_{\theta,N}^*(\lambda) = 2\frac{1}{N} \sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]} \widetilde{I}_{x^*}^T(\lambda_j) / |\Psi_\theta(\lambda_j)|^2$.

Other procedures are possible as that based on that of Franke and Härdle (1992), where the bootstrap periodogram $\widetilde{I}_{x^*}^{T}(\lambda_j) = \left|\Psi_{\widehat{\theta}}(\lambda_j)\right|^2 \check{I}_{x^*}^{T}(\lambda_j)$ is replaced by

$$\mathring{I}_{x^*}^{T}(\lambda_j) = f_{\widehat{\vartheta}}(\lambda_j)\,\chi_j,$$

where $\chi_{-\tilde{n}}, ..., \chi_{\tilde{n}}$ are independent exponential random variables. However, unlike in the case of $d = 1$, the previous bootstrap algorithm will not be valid. The reason is because the bootstrap does not correctly "estimate" the fourth cumulant κ_ε. More specifically the asymptotic distribution of the bootstrap estimator $\check{\vartheta}^*$ in (4.2) will not have the same asymptotic variance as that of $\widehat{\vartheta}$ in (3.1).

So to overcome this problem, following Hidalgo (2007), see also Hidalgo and Lazarova (2007), we propose in the paper an alternative algorithm, as described in the next *4 STEPS.*

STEP 1: We first obtain the residuals

$$\widehat{\varepsilon}(t) = (2\pi)^{d/2} \frac{1}{N^{1/2}} \sum_{j=-\tilde{n}}^{\tilde{n}} e^{-it\cdot\lambda_j} \Psi_{\widehat{\theta}}^{-1}(\lambda_j)\, w_x(\lambda_j),$$

for $t = 1, ..., n$. From here as usual, we obtain a random sample of size N with replacement from the empirical distribution function of $\{\widehat{\varepsilon}(t)\}_{t=1}^{n}$. Let's denote the bootstrap sample by $\{\varepsilon^*(t)\}_{t=1}^{n}$.

Remark 4. (**a**) *Notice that because $\widehat{\eta}_N = \eta\left(\alpha_{\widehat{\theta},N}\right)$ is asymptotically independent of the mean and variance of $\{\varepsilon(t)\}_{t\in\mathbb{Z}^d}$, we do not need to standardize $\widehat{\varepsilon}(t)$ to obtain the bootstrap sample.* (**b**) *The motivation to compute the residuals as in* STEP 1 *comes from the observation that, for $t = 1, ..., n$,*

$$\varepsilon(t) \simeq (2\pi)^{d/2} \frac{1}{N^{1/2}} \sum_{j=-\tilde{n}}^{\tilde{n}} e^{-it\cdot\lambda_j} \Psi_{\theta_0}^{-1}(\lambda_j)\, w_x(\lambda_j).$$

STEP 2: For $t = 1, ..., n$, compute the bootstrap observations

$$x^*(t) = (2\pi)^{d/2} \frac{1}{N^{1/2}} \sum_{j=-\tilde{n}}^{\tilde{n}} e^{-it\cdot\lambda_j} \Psi_{\widehat{\theta}}(\lambda_j)\, w_{\varepsilon^*}(\lambda_j), \tag{4.3}$$

where $w_{\varepsilon^*}(\lambda_j)$ is the standard *DFT* of $\{\varepsilon^*(t)\}_{t=1}^{n}$, and the taper periodogram $I_{x^*}^{T}(\lambda_j)$ as defined in (2.4).

STEP 3: The bootstrap analogue of $\widehat{\vartheta}$ is given by

$$\widehat{\vartheta}^* = \arg\min_{\vartheta\in\Theta\times\mathbb{R}^+} \mathcal{Q}_N^*(\vartheta), \tag{4.4}$$

where

$$\mathcal{Q}_N^*(\vartheta) = \frac{1}{N} \sum_{j=-\tilde{n}}^{\tilde{n}} \left\{ \log f_\vartheta(\lambda_j) + \frac{I_{x^*}^{T}(\lambda_j)}{(2\pi)^d f_\vartheta(\lambda_j)} \right\}. \tag{4.5}$$

STEP 4: Compute the bootstrap $T_p - process$

$$\alpha_{\widehat{\theta}^*,N}^{*}(\lambda) = 2^{-1/2} N^{1/2} \left[\frac{G_{\widehat{\theta}^*,N}^{*}(\lambda)}{G_{\widehat{\theta}^*,N}^{*}(\pi)} - \prod_{\ell=1}^{d} \left(\frac{\lambda[\ell]}{\pi} \right) \right], \quad \lambda \in [0,\pi]^d, \tag{4.6}$$

with $G_{\theta,N}^{*}(\lambda) = 2N^{-1} \sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]} \left|\Psi_\theta(\lambda_j)\right|^{-2} I_{x^*}^{T}(\lambda_j)$.

Theorem 4. *Under $C1 - C3$ and $C5 - C8$, we have that*

$$N^{1/2}\left(\widehat{\vartheta}^* - \widehat{\vartheta}\right) \xrightarrow{d^*} \mathcal{N}\left(0, 2\Lambda_{\vartheta_0}^{-1} V_{\vartheta_0} \Lambda_{\vartheta_0}^{-1}\right),$$

where $\xrightarrow{d^*}$ denotes convergence in bootstrap distribution.

As with $\widehat{\vartheta}$ in Section 3, $\widehat{\vartheta}^*$ in (4.4) satisfies the asymptotic linearization

$$\widehat{\vartheta}^* - \widehat{\vartheta} = -Q_{\widehat{\vartheta},N}^{-1} \left\{ \int_{-\pi}^{\pi} \widetilde{\phi}_{\widehat{\theta}}(\lambda)\, \alpha_{\widehat{\theta},N}(d\lambda) + \int_{-\pi}^{\pi} \phi_{\widehat{\vartheta}}(\lambda)\, d\lambda \frac{1}{N} \sum_{j=-\tilde{n}}^{\tilde{n}} \left(\frac{I_{x^*}^T(\lambda_j)}{(2\pi)^d f_{\widehat{\vartheta}}(\lambda_j)} - 1 \right) \right\} \tag{4.7}$$
$$+ o_{p^*}\left(N^{-1/2}\right).$$

Denote $G_N^{*0}(\lambda) = 2\frac{1}{N}\sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]} I_{\varepsilon^*}^T(\lambda_j)$ and let

$$\alpha_N^{*0}(\lambda) = 2^{-1/2}N^{1/2}\left[\frac{G_N^{*0}(\lambda)}{G_N^{*0}(\pi)} - \prod_{\ell=1}^{d}\left(\frac{\lambda[\ell]}{\pi}\right)\right], \quad \lambda \in [0,\pi]^d. \tag{4.8}$$

Theorem 5. *Under H_0 and assuming $C.1 - C.3$ and $C5 - C8$, uniformly in $\lambda \in [0,\pi]$,*

$$(a)\quad \alpha_{\widehat{\theta}^*,N}^*(\lambda) = \alpha_N^{*0}(\lambda) - \left(\frac{1}{N}\sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]} \widetilde{\varphi}_{\widehat{\theta},N}'(\lambda_j)\right) \widetilde{\Lambda}_{\widehat{\theta},N}^{-1} \frac{1}{N}\sum_{j=-\tilde{n}}^{\tilde{n}} \widetilde{\varphi}_{\widehat{\theta},N}'(\lambda_j)\, I_{\varepsilon^*}^T(\lambda_j)$$
$$+ o_{p^*}(1).$$

$$(b)\quad \alpha_{\widehat{\theta}^*,N}^* \overset{d^*}{\Rightarrow} \alpha_\infty.$$

A conclusion from Theorem 5 is the following corollary.

Corollary 3. *Under the maintained hypothesis and assuming $C1 - C3$ and $C5 - C8$, we have that for any continuous functional η,*

$$\widehat{\eta}_N^* := \eta\left(\alpha_{\widehat{\theta}^*,N}^*\right) \overset{d^*}{\to} \eta(\alpha_\infty).$$

Proof. The proof follows from Theorem 5 and the continuous mapping theorem. □

Thus, Theorem 5 and Corollary 3 indicate that the bootstrap statistic $\widehat{\eta}_N^*$ is consistent. That is, let $c_{N,(1-\alpha)}^f$ and $c_{(1-\alpha)}^a$ be such that

$$\Pr\left\{|\widehat{\eta}_N| > c_{n,(1-\alpha)}^f\right\} = \alpha; \qquad \lim_{n\to\infty}\Pr\left\{|\widehat{\eta}_N| > c_{(1-\alpha)}^a\right\} = \alpha,$$

respectively. So, Theorems 3 and 5 indicate that $c_{N,(1-\alpha)}^f \to c_{(1-\alpha)}^a$ and $c_{(1-\alpha)}^* \overset{p}{\to} c_{(1-\alpha)}^a$, respectively, where $c_{(1-\alpha)}^*$ is defined as $\Pr\left\{|\widehat{\eta}_N^*| > c_{(1-\alpha)}^*\right\} = \alpha$.

Typically, the finite sample distribution of $\widehat{\eta}_N^*$ is not available, although the critical values $c_{(1-\alpha)}^*$ can be approximated, as accurately as desired, by standard Monte-Carlo simulation. To that end, consider the bootstrap samples $\left\{\widehat{\varepsilon}^{*\ell}(t)\right\}_{t=1}^n$ for $\ell = 1, ..., B$, and compute $\alpha_{\widehat{\theta}^*,N}^\ell(\lambda)$ as in (4.5) for each ℓ. Then, $c_{(1-\alpha)}^*$ is approximated by the value $c_{(1-\alpha)}^{*B}$ that satisfies $B^{-1}\sum_{\ell=1}^B \mathcal{I}\left(\eta\left(\alpha_{\widehat{\theta}^*,N}^\ell\right) > c_{(1-\alpha)}^{*B}\right) = \alpha$.

Next we study the behaviour of the bootstrap tests under the alternative hypothesis.

Corollary 4. *Assuming C.1-C.8, under H_1,*

$$\widehat{\eta}_N^* \overset{d^*}{\to} \eta(\widetilde{\alpha}_\infty) \ \textit{in probability},$$

where $\widetilde{\alpha}_\infty$ is a centered Gaussian process with covariance structure as α_∞ but with θ_0 replaced by $\theta_1 = \text{plim}\,\widehat{\theta}$.

Proof. The proof proceeds exactly as that of Theorem 5 and then Corollary 3 but instead of writing $\widehat{\theta} - \theta_0 = o_p(1)$ we write $\widehat{\theta} - \theta_1 = o_p(1)$ and θ_1 instead of θ_0. □

5. LEMMAS

First, we introduce some notation. We denote the conjugate of a complex number a by $\overline{a}$. Also, for a generic function $\nu(\lambda)$, we abbreviate $\nu(\lambda_j)$ by $\nu_j = \left(\nu_{j[1]}, ..., \nu_{j[d]}\right)'$ and C will denote a generic positive and finite constant.

For the next two lemmas, we shall assume that $\{\xi(t)\}_{t\in\mathbb{Z}^d}$ and $\{\zeta(t)\}_{t\in\mathbb{Z}^d}$ are two stationary spatial processes with a representation as that in (1.1) and whose respective errors satisfy $C1$. Also $f_{\xi\zeta}(\lambda) = (2\pi)^{-d} \sum_{j\in\mathbb{Z}^d} \mathbb{E}(\xi(t)\zeta(t+j)) \exp\{-ij\cdot\lambda\}$, the cross-spectral density function, is a twice continuously differentiable function in $\lambda \in \Pi^d$. Denote $\widetilde{\mathbb{Z}}^d = \{j : (-\tilde{n} \prec j \prec \tilde{n}) \wedge (0 < j[1])\}$.

Lemma 1. *Consider $j \in \widetilde{\mathbb{Z}}^d$. Then,*

$$(a)\ \mathbb{E}\left(w^T_{\xi,j}\overline{w}^T_{\zeta,j}\right) - f_{\xi\zeta,j} = O\left(\vec{n}^{-2}\right); \quad (b)\ \mathbb{E}\left(w^T_{\xi,j}w^T_{\zeta,j}\right) = O\left(\prod_{\ell=1}^{d} j[\ell]^{-3}\right).$$

Proof. We begin with part (a). By definition, the left side of the equality in (a) is

$$\int_{\Pi^{d\vec{n}}} \left(f_{\xi\zeta}(\lambda) - f_{\xi\zeta}(\lambda_j)\right) \prod_{\ell=1}^{d} K^T\left(\lambda[\ell] - \lambda_{j[\ell]}\right) d\lambda,$$

suppressing any reference to ℓ in K^T_ℓ and/or D^T_ℓ for notational simplicity.

Now, because $f_{\xi\zeta}(\lambda)$ is twice continuously differentiable and $\int_\Pi \mu K^T(\mu)\, d\mu = 0$, we have that the last displayed expression is bounded in modulus by

$$\begin{aligned} & C\int_{\Pi^d} \sum_{\ell=1}^{d}\sum_{p=1}^{d} \left|\lambda[\ell] - \lambda_{j[\ell]}\right| \left|\lambda[p] - \lambda_{j[p]}\right| \prod_{\ell=1}^{d} K^T\left(\lambda[\ell] - \lambda_{j[\ell]}\right) d\lambda \\ \leq\ & C\int_{\Pi^d} \sum_{\ell=1}^{d} \left|\lambda[\ell] - \lambda_{j[\ell]}\right|^2 \prod_{\ell=1}^{d} K^T\left(\lambda[\ell] - \lambda_{j[\ell]}\right) d\lambda \end{aligned}$$

by the Cauchy-Schwarz inequality. Now, using (2.6), that the Fejer's kernel integrates 1 and that $\sum_{t[\ell]=1}^{n[\ell]} h_\ell(t[\ell])^2 \geq C\vec{n}$, we obtain that the right side of the last displayed inequality is, by $C2$ and standard algebra, bounded by

$$\frac{C}{N}\int_{\Pi^d} \sum_{\ell=1}^{d} \left|\lambda[\ell] - \lambda_{j[\ell]}\right|^2 \prod_{\ell=1}^{d} \min\left\{\vec{n}^2, \vec{n}^{-4}\left|\lambda[\ell] - \lambda_{j[\ell]}\right|^{-6}\right\} d\lambda = O\left(\vec{n}^{-2}\right).$$

Next we show part (b). Again by definition and that $|f_{\xi\zeta}(\lambda)| < C$, we obtain that $\left|\mathbb{E}\left(w^T_{\xi,j}w^T_{\zeta,j}\right)\right|$ is bounded by

$$C\int_{\Pi^d} |f_{\xi\zeta}(\lambda)| \prod_{\ell=1}^{d} n^{-1}\left|D^T\left(\lambda[\ell] - \lambda_{j[\ell]}\right) D^T\left(\lambda[\ell] + \lambda_{j[\ell]}\right)\right| d\lambda \leq C\prod_{\ell=1}^{d} j[\ell]^{-3}$$

by standard arguments after using (2.6). □

Lemma 2. *Let $k \prec j \in \widetilde{\mathbb{Z}}^d$ and $c_{jk} = \min\left\{\prod_{\ell=1}^{d} |j[\ell] - k[\ell]|_+^{-3}, \frac{\log \vec{n}}{\vec{n}}\right\}$. Then,*

$$\begin{aligned} (a)\quad \mathbb{E}\left(w^T_{\xi,j}\overline{w}^T_{\zeta,k}\right) &= f_{\xi\zeta,j}\mathcal{I}\left(|j[\ell] - k[\ell]| = 2, \ell = 1, ..., d\right) + O(c_{jk}) \\ (b)\quad \mathbb{E}\left(w^T_{\xi,j}w^T_{\zeta,k}\right) &= O(c_{jk}). \end{aligned}$$

Proof. We shall handle part (a) only, being part (b) identical. By definition, (5.1)

$$\mathbb{E}\left(w^T_{\xi,j}\overline{w}^T_{\zeta,k}\right) = 6^{-d}\int_{\Pi^d} f_{\xi\zeta}(\lambda) \prod_{\ell=1}^{d} n[\ell]^{-1} D^T\left(\lambda[\ell] - \lambda_{j[\ell]}\right) D^T\left(\lambda_{k[\ell]} - \lambda[\ell]\right) d\lambda.$$

Because $|f_{\xi\zeta}(\lambda)| < C$, the modulus of the right side of (5.1) is bounded by

$$C \prod_{\ell=1}^{d} \vec{n}^{-1} \left\{ \int_{-\pi}^{\lambda_{(j[\ell]+k[\ell])/2}} + \int_{\lambda_{(j[\ell]+k[\ell])/2}}^{\pi} \right\} \left| D^T\left(\lambda[\ell] - \lambda_{j[\ell]}\right) \right| \left| D^T\left(\lambda_{k[\ell]} - \lambda[\ell]\right) \right| d\lambda[\ell]$$

using $C2$. Now using (2.6) and because $\int_0^\pi \left| D^T(\lambda) \right| d\lambda < C$, the contribution due to a factor of the type $\int_{-\pi}^{\lambda_{(j[\ell]+k[\ell])/2}}$ when $\lambda_{k[\ell]} < \lambda_{j[\ell]}$ is bounded by

$$C\, |j[\ell] - k[\ell]|_+^{-3} \int_{-\pi}^{\lambda_{(j[\ell]+k[\ell])/2}} \left| D^T\left(\lambda_{k[\ell]} - \lambda[\ell]\right) \right| d\lambda[\ell] = O\left(|j[\ell] - k[\ell]|_+^{-3} \right),$$

whereas if $\lambda_{j[\ell]} < \lambda_{k[\ell]}$ by $C\,|j[\ell] - k[\ell]|_+^{-3} \int_{-\pi}^{\lambda_{(j[\ell]+k[\ell])/2}} \left| D^T\left(\lambda[\ell] - \lambda_{j[\ell]}\right) \right| d\lambda[\ell] = O\left(|j[\ell] - k[\ell]|_+^{-3} \right)$. Recall that $k \prec j$ so we have for some $\ell = 1, .., d-1$, $\lambda_{j[\ell]} < \lambda_{k[\ell]}$. Finally, proceeding similarly the contribution due to a factor of the type $\int_{\lambda_{(j[\ell]+k[\ell])/2}}^{\pi}$ is $O\left(|j[\ell] - k[\ell]|_+^{-3} \right)$. Now conclude by Hölder's inequality and that $\left| \mathbb{E}\left(w_{\xi,j}^T \overline{w}_{\zeta,k}^T \right) \right| = O\left(\prod_{\ell=1}^{d} |j[\ell] - k[\ell]|_+^{-3} \right)$. On the other hand, because when $|j[\ell] - k[\ell]| \neq 2$ for some $\ell = 1, ..., d$,

$$\int_{\Pi^d} \prod_{\ell=1}^{d} D^T\left(\lambda[\ell] - \lambda_{j[\ell]}\right) D^T\left(\lambda_{k[\ell]} - \lambda[\ell]\right) d\lambda = 0, \tag{5.2}$$

we have that in this case, except multiplicative constants, the left side of (5.1) is

$$N^{-1} \int_{\Pi^d} \left(f_{\xi\zeta}(\lambda) - f_{\xi\zeta}(\lambda_j) \right) \prod_{\ell=1}^{d} D\left(\lambda[\ell] - \lambda_{j[\ell]}\right) D\left(\lambda_{k[\ell]} - \lambda[\ell]\right) d\lambda,$$

which, by the mean value theorem, is bounded in absolute value by

$$N^{-1} \int_{\Pi^d} \sum_{\ell=1}^{d} \left| \lambda[\ell] - \lambda_{j[\ell]} \right| \prod_{\ell=1}^{d} \left| D\left(\lambda[\ell] - \lambda_{j[\ell]}\right) \right| \left| D\left(\lambda_{k[\ell]} - \lambda[\ell]\right) \right| d\lambda = O\left(\frac{\log \vec{n}}{\vec{n}} \right),$$

because $|\lambda D(\lambda)| < C$, $\int_0^\pi |D(\lambda)|\, d\lambda = O(\log \vec{n})$ and the Cauchy-Schwarz inequality imply that $\int_\Pi \left| D\left(\lambda - \lambda_{j[\ell]}\right) \right| \left| D\left(\lambda_{k[\ell]} - \lambda\right) \right| d\lambda = O(\vec{n})$. Now, when for all $\ell = 1, ..., d$, $|j[\ell] - k[\ell]| = 2$, because the left side of (5.2) is 1, we have that proceeding as above, the left side of (5.1) is $f_{\xi\zeta,j} + O\left(\vec{n}^{-1} \log \vec{n} \right)$. This concludes the proof. $\square$

In what follows, we shall abbreviate $w_x^T(\lambda) / \Psi(\lambda)$ and $w_\varepsilon^T(\lambda)$ by $u(\lambda)$ and $v(\lambda)$ respectively for all $\lambda \in \Pi^d$.

Lemma 3. *Let $\zeta(\lambda)$ be a continuously differentiable function in Π^d. Under C1-C4, we have that for all $r \leq s \in \widetilde{\mathbb{Z}}^d$*

$$\mathbb{E} \left| \sum_{j=r}^{s} \zeta_j v_j \left(u_j - v_j \right) \right|^2 = O\left(\frac{\log \vec{n}}{\vec{n}} \prod_{\ell=1}^{d} |s[\ell] - r[\ell]|_+ \right). \tag{5.3}$$

Proof. Denote $\varrho_j = u_j - v_j$. By standard arguments, the left side of (5.3) is

$$\begin{aligned} & \sum_{j=r}^{s} \zeta_j^2 \mathbb{E}\left\{ v_j \overline{v}_j \overline{\varrho}_j \varrho_j \right\} + \sum_{j \neq k = r}^{s} \zeta_j \zeta_k \mathbb{E}\left\{ v_j \overline{v}_k \varrho_j \overline{\varrho}_k \right\} \\ = & \sum_{j=r}^{s} \zeta_j^2 \left\{ a_{j1} + a_{j2} \right\} + \sum_{j \neq k = r}^{s} \zeta_j \zeta_k \left\{ b_{jk,1} + b_{jk,2} \right\}, \end{aligned}$$

where

$$
\begin{aligned}
a_{j1} &= \mathbb{E}\left(v_j\overline{v}_j\right)\mathbb{E}\left(\varrho_j\overline{\varrho}_j\right) + \left|\mathbb{E}\left(v_j\overline{\varrho}_j\right)\right|^2 + \left|\mathbb{E}\left(v_j\varrho_j\right)\right|^2 \\
a_{j2} &= cum\left(v_j,\overline{v}_j,\overline{u}_j,u_j\right) + cum\left(v_j,\overline{v}_j,\overline{v}_j,v_j\right) \\
&\quad - cum\left(v_j,\overline{v}_j,\overline{u}_j,v_j\right) - cum\left(v_j,\overline{v}_j,u_j,\overline{v}_j\right) \\
b_{jk,1} &= \mathbb{E}\left(v_j\overline{v}_k\right)\mathbb{E}\left(\varrho_j\overline{\varrho}_k\right) + \mathbb{E}\left(v_j\varrho_j\right)\mathbb{E}\left(\overline{v}_k\overline{\varrho}_k\right) + \mathbb{E}\left(v_j\overline{\varrho}_k\right)\mathbb{E}\left(\overline{v}_k\varrho_j\right) \\
b_{jk,2} &= cum\left(v_j,\overline{v}_k,\overline{u}_j,u_k\right) + cum\left(v_j,\overline{v}_k,\overline{v}_j,v_k\right) \\
&\quad - cum\left(v_j,\overline{v}_k,\overline{u}_j,v_k\right) - cum\left(v_j,\overline{v}_k,u_j,\overline{v}_k\right).
\end{aligned}
$$

After observing that $\mathbb{E}\left(v_j\overline{u}_j\right) = 1 + O\left(\vec{n}^{-2}\right)$ and $\mathbb{E}\left(v_j\overline{\varrho}_j\right) = \mathbb{E}\left(v_j\overline{u}_j\right) - \mathbb{E}\left(v_j\overline{v}_j\right)$, we have that Lemma 1 implies that $a_{j1} = O\left(\vec{n}^{-2}\right)$, whereas Lemmas 1 and 2 imply that $b_{jk,1} = O\left(c_{jk}^2 + \vec{n}^{-1}\log n\mathcal{I}\left(\left|j\left[\ell\right] - k\left[\ell\right]\right| = 2, \ell = 1,...,d\right)\right)$, with c_{jk} as defined there. From here it is immediate to conclude that the contribution due to a_{j1} and $b_{jk,1}$ into the left of (5.3) is its right side.

Finally we examine a_{j2} and $b_{jk,2}$. Using formulae in Brillinger [(1981), (2.6.3), page 26, and (2.10.3), page 39], we deduce after standard algebra that

$$
\begin{aligned}
b_{jk,2} &= \frac{\kappa_\varepsilon}{N^2}\int_{\Pi^d}\int_{\Pi^d}\left(\frac{\Psi\left(\lambda\right)}{\Psi_j} - 1\right)\left(\frac{\Psi\left(\mu\right)}{\Psi_k} - 1\right)D^T\left(\lambda - \lambda_j\right)D^T\left(\mu + \lambda_k\right) \\
&\quad \times D^T\left(\lambda_j - \lambda_k - \lambda - \mu\right)d\lambda d\mu.
\end{aligned}
$$

By the Cauchy-Schwarz inequality, we have that $\left|b_{jk,2}\right|^2$ is bounded by CN^{-1} times

$$
\int_{\Pi^d}\left(\frac{\Psi\left(\lambda\right)}{\Psi_j} - 1\right)^2 K^T\left(\lambda - \lambda_j\right)d\lambda\int_{\Pi^d}\left(\frac{\Psi\left(\mu\right)}{\Psi_k} - 1\right)^2 K^T\left(\mu + \lambda_k\right)K^T\left(\lambda_j - \lambda_k - \lambda - \mu\right)d\lambda d\mu.
$$

Proceeding as in Lemma 2 and by $C4$, we then obtain that $b_{jk,2} = O\left(\vec{n}^{-2}N^{-1/2}\right)$. Likewise $a_{j2} = O\left(\vec{n}^{-2}N^{-1/2}\right)$. From here, the conclusion of the lemma easily follows by observing that $\prod_{\ell=1}^{d}\left|s\left[\ell\right] - r\left[\ell\right]\right|_+ \leq N$. □

Lemma 4. *Let $\zeta\left(\lambda\right)$ be a function as in Lemma 3. Then, under $C1 - C4$,*

$$
\mathbb{E}\sup_{\lambda\in[0,\pi]^d}\left|\sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]}\zeta_j\left\{\frac{I_{x,j}^T}{\left|\Psi_j\right|^2} - I_{\varepsilon,j}^T\right\}\right| = o\left(N^{1/2}\right).
$$

Proof. We shall consider the proof in the positive quadrant $\sum_{j=1}^{[\tilde{n}\lambda/\pi]}$, being the proof for the remaining $2^{d-1} - 1$ quadrants similarly handled. By the Cauchy-Schwarz and the triangle inequalities, it suffices to show that

$$
\mathbb{E}\sup_s\left|\sum_{j=1}^{s}\zeta_j\left\{\frac{I_{x,j}^T}{\left|\Psi_j\right|^2} - I_{\varepsilon,j}^T\right\}\right| \leq \mathbb{E}\sup_s\sum_{j=1}^{s}\left|\zeta_j\right|\left|\varrho_j\right|^2 + 2\mathbb{E}\sup_s\left|\sum_{j=1}^{s}\zeta_j v_j\overline{\varrho}_j\right| \tag{5.4}
$$

is $o\left(N^{1/2}\right)$, where we abbreviate "$\sup_{s=1,...,\tilde{n}}$" by "$\sup_s$" and $\varrho_j = u_j - v_j$.

The first term on the right of (5.4) is bounded by

$$
C\sum_{j=1}^{\tilde{n}}\left\{\left(\mathbb{E}\left|u_j\right|^2 - 1\right) - \left(\mathbb{E}\left(u_j\overline{v}_j\right) - 1\right) - \left(\mathbb{E}\left(\overline{u}_j v_j\right) - 1\right) + \left(\mathbb{E}\left|v_j\right|^2 - 1\right)\right\} = o\left(N^{1/2}\right),
$$

because $\left|\zeta_j\right| \leq C$, $d < 4$ and by Lemma 1, for instance

$$
\left|\mathbb{E}\left(u_j\left\{\begin{array}{c}\overline{v}_j\\ \overline{u}_j\end{array}\right)-1\right| \leq \sigma_\varepsilon^{-2}\left|\Psi_j\right|^{-2}\left(\left|\mathbb{E}\left(w_{x,j}^T\left\{\begin{array}{c}\overline{w}_{\varepsilon,j}^T\\ \overline{w}_{x,j}^T\end{array}\right) - \sigma_\varepsilon^2\left\{\begin{array}{c}\Psi_j\\ \left|\Psi_j\right|^2\end{array}\right|\right) = O\left(\frac{1}{\vec{n}^2}\right).
$$

Next, we examine the second term of (5.4). To that end, let $q = 0,\ldots,\left[\tilde{n}^\varsigma\right] - 1$ for some $0 < \varsigma < 1/d$. (Recall that $\left[\tilde{n}^\psi\right] = \left(\left[\tilde{n}^\psi\left[1\right]\right],...\left[\tilde{n}^\psi\left[d\right]\right]\right)$ for any $\psi > 0$.)

Standard inequalities imply that the square of the second term on the right of (5.4) is bounded by

$$\mathbb{E}\max_{s}\left|\left\{\sum_{j=1}^{s}-\sum_{j=1}^{q(s)\left[\tilde{n}^{1-\varsigma}\right]}\right\}\zeta_j v_j \overline{\varrho}_j\right|^2+\mathbb{E}\max_{s}\left|\sum_{j=1}^{q(s)\left[\tilde{n}^{1-\varsigma}\right]}\zeta_j v_j \overline{\varrho}_j\right|^2, \tag{5.5}$$

where herewith $q(s)$ denotes the value of $q=0,\ldots,[\tilde{n}^{\varsigma}]-1$ such that $q(s)\left[\tilde{n}^{1-\varsigma}\right]$ is the largest vector s_1 such that $s_1\le s$, and using the convention $\sum_{j=c}^{d}\equiv 0$ if $d<c$. From now on, we abbreviate $(\tilde{n}[1]/[\tilde{n}^{\varsigma}[1]],...,\tilde{n}[d]/[\tilde{n}^{\varsigma}[d]])$ by $\left[\tilde{n}^{1-\varsigma}\right]$.

From the definition of $q(s)$ and $\left(\sup_p|c_p|\right)^2=\sup_p|c_p|^2\le\sum_p|c_p|^2$, the second term of (5.5) is bounded by $\sum_{q=1}^{[\tilde{n}^{\varsigma}]-1}\mathbb{E}\left|\sum_{j=1}^{q\left[\tilde{n}^{1-\varsigma}\right]}\zeta_j v_j\overline{\varrho}_j\right|^2=O\left(N^{1+\varsigma}\vec{n}^{-1}\log^2\vec{n}\right)=o(N)$ by Lemma 3 and because $\varsigma<1/d$.

To complete the proof we need to show that the first term in (5.5) is $o(N)$. To that end, we note that it is bounded by

$$\mathbb{E}\max_{q=1,\ldots,[\tilde{n}^{\varsigma}]-1}\ \max_{s=1+q[\tilde{n}^{1-\varsigma}],\ldots,(q+1)[\tilde{n}^{1-\varsigma}]}\left|\sum_{j=1+q[\tilde{n}^{1-\varsigma}]}^{s}\zeta_j v_j\overline{\varrho}_j\right|^2$$

which is $O(N^{\varsigma})\,\mathbb{E}\max_{s=1,\ldots,[\tilde{n}^{1-\varsigma}]}\left|\sum_{j=1}^{s}\zeta_j v_j\overline{\varrho}_j\right|^2$.

So, we have that the square of the second term on the right of (5.4) is

$$o(N)+O(N^{\varsigma})\,\mathbb{E}\max_{s=1,\ldots,[\tilde{n}^{1-\varsigma}]}\left|\sum_{j=1}^{s}\zeta_j v_j\overline{\varrho}_j\right|^2.$$

Observe that the second factor of the second term of the last displayed expression is similar to the second term on the right of (5.4) but with $s=1,...,\left[\tilde{n}^{1-\varsigma}\right]$ instead of $s=1,...,\tilde{n}$. So, repeating the same steps, the last displayed expression, and so the square of the second term on the right of (5.4), is

$$\begin{aligned}&o(N)+N^{\varsigma\sum_{p=0}^{\iota-1}(1-\varsigma)^p}\,\mathbb{E}\max_{s=1,\ldots,([\tilde{n}^{1-\varsigma}])^{\iota}}\left|\sum_{j=1}^{s}\zeta_j v_j\overline{\varrho}_j\right|^2\\ =\ &o(N)+O\left(N^{\varsigma\sum_{p=0}^{\iota-1}(1-\varsigma)^p}\right)\sum_{s=1}^{\tilde{n}^{(1-\varsigma)^{\iota}}}\mathbb{E}\left|\sum_{j=1}^{s}\zeta_j v_j\overline{\varrho}_j\right|^2=o(N)\end{aligned}$$

after choosing ι large enough because $\varsigma<1/d$. This completes the proof. □

Lemma 5. *Let $\zeta(\lambda;\vartheta)$ be as in Lemma 3 for all $\vartheta\in\Theta\times\mathbb{R}^+$, and continuously differentiable in ϑ for all λ. Assuming $C1-C4$,*

$$\frac{1}{N}\sup_{\vartheta\in\Theta\times\mathbb{R}^+}\left|\sum_{j=-\tilde{n}}^{\tilde{n}}\zeta_j(\vartheta)\left(\frac{I_{x,j}^T}{\left|\Psi_{\theta_0,j}\right|^2}-1\right)\right|=o_p(1). \tag{5.6}$$

Proof. By the triangle inequality, the left side of (5.6) is bounded by

$$\frac{C}{N}\sup_{\vartheta\in\Theta\times\mathbb{R}^+}\left|\sum_{j=-\tilde{n}}^{\tilde{n}}\zeta_j(\vartheta)\left(\frac{I_{x,j}^T}{\left|\Psi_{\theta_0,j}\right|^2}-I_{\varepsilon,j}^T\right)\right|+\frac{C}{N}\sup_{\vartheta\in\Theta\times\mathbb{R}^+}\left|\sum_{j=-\tilde{n}}^{\tilde{n}}\zeta_j(\vartheta)\left(I_{\varepsilon,j}^T-1\right)\right|. \tag{5.7}$$

Now, because by assumption $|\zeta(\lambda;\vartheta)| < C$, the first term of (5.7) is bounded by

$$\frac{C}{N}\sum_{j=-\tilde{n}}^{\tilde{n}}\left|\frac{I^T_{x,j}}{\left|\Psi_{\theta_0,j}\right|^2} - I^T_{\varepsilon,j}\right| \leq \frac{C}{N}\sum_{j=-\tilde{n}}^{\tilde{n}}\left|\varrho_j\right|^2 + \frac{C}{N}\sum_{j=-\tilde{n}}^{\tilde{n}}\left|v_j\overline{\varrho}_j\right| = o_p(1)$$

by Markov's inequality because by the Cauchy-Schwarz inequality $\mathbb{E}\left|v_j\overline{\varrho}_j\right|^2 \leq \mathbb{E}\left|v_j\right|^2\mathbb{E}\left|\overline{\varrho}_j\right|^2$ and then proceeding as in Lemma 3. Next, we show that the second term of (5.7) is $o_p(1)$. But this follows by standard arguments (see also Lemma 15) and because $\zeta_j(\vartheta)$ is continuously differentiable in ϑ. □

Lemma 6. *Assume $C1-C3$ and $C5-C8$. Then, $\widehat{\vartheta} - \vartheta_0 = o_p(1)$.*

Proof. The proof follows very easily using Lemma 5. Indeed, (3.2) is

$$\frac{1}{N}\sum_{j=-\tilde{n}}^{\tilde{n}}\frac{f_{\vartheta_0,j}}{f_{\vartheta,j}}\left(\frac{I^T_{x,j}}{(2\pi)^d f_{\vartheta_0,j}} - 1\right) + \frac{1}{N}\left\{\sum_{j=-\tilde{n}}^{\tilde{n}}\frac{f_{\vartheta_0,j}}{f_{\vartheta,j}} - \log\frac{f_{\vartheta_0,j}}{f_{\vartheta,j}} + \log f_{\vartheta_0,j}\right\}. \tag{5.8}$$

Now the second term of (5.8) converges using Brillinger (1981, *p*.15) to

$$\int_{\pi}^{\pi}\left\{\frac{f_{\vartheta_0}(\lambda)}{f_{\vartheta}(\lambda)} - \log\left(\frac{f_{\vartheta_0}(\lambda)}{f_{\vartheta}(\lambda)}\right)\right\}d\lambda + \int_{\pi}^{\pi}\log f_{\vartheta_0}(\lambda)\,d\lambda \geq \frac{(2\pi)^d}{2} + \int_{\pi}^{\pi}\log f_{\vartheta_0}(\lambda)\,d\lambda$$

with equality when $f_{\vartheta_0}(\lambda) = f_{\vartheta}(\lambda)$ which is the case only if $\vartheta = \vartheta_0$ by $C7$. On the other hand, the first term of (5.8) converges to zero uniformly in ϑ by Lemma 5 because $f^{-1}_{\vartheta,j}f_{\vartheta_0,j}$ satisfies the same conditions as $\zeta(\lambda;\vartheta)$ there by $C6$. From here the conclusion of the lemma is standard, so we omit its details. □

Lemma 7. *Assume $C1-C3$ and $C5-C8$. Under H_0, uniform in $\lambda \in [0,\pi]^d$,*

$$\frac{1}{N^{1/2}}\sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]}\zeta_j\left(\frac{I^T_{x,j}}{\left|\Psi_{\widehat{\theta},j}\right|^2} - I^T_{\varepsilon,j}\right) = -\left(\frac{1}{N}\sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]}\zeta_j\varphi'_{\theta_0,j}\right)N^{1/2}\left(\widehat{\theta} - \theta_0\right) + o_p(1), \tag{5.9}$$

where $\zeta(\lambda)$ is as in Lemma 3.

Proof. The difference between the left side of (5.9) and the first term on its right side is

$$\begin{aligned}&\frac{1}{N^{1/2}}\sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]}\zeta_j\frac{I^T_{x,j}}{\left|\Psi_{\theta_0,j}\right|^2}\left[\frac{\left|\Psi_{\theta_0,j}\right|^2}{\left|\Psi_{\widehat{\theta},j}\right|^2} - 1 + \varphi'_{\theta_0,j}\left(\widehat{\theta} - \theta_0\right)\right] \\ &+ \frac{1}{N^{1/2}}\sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]}\zeta_j\left(\frac{I^T_{x,j}}{\left|\Psi_{\theta_0,j}\right|^2} - I^T_{\varepsilon,j}\right) - \frac{1}{N^{1/2}}\sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]}\zeta_j\varphi'_{\theta_0,j}\frac{I^T_{x,j}}{\left|\Psi_{\theta_0,j}\right|^2}\left(\widehat{\theta} - \theta_0\right).\end{aligned} \tag{5.10}$$

First, because each component of the vector $\zeta(\lambda)\varphi_{\theta_0}(\lambda)$ satisfies the same conditions of $\zeta(\lambda)$ in Lemma 4, Markov's inequality implies that the second term of (5.10) is $o_p(1)$, whereas the third term is $N^{-1}\sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]}\zeta_j\varphi'_{\theta_0,j}N^{1/2}\left(\widehat{\theta} - \theta_0\right) + o_p(1)$ by Lemma 4 and because proceeding as in Robinson and Vidal-Sanz (2006)

$$\frac{1}{N^{1/2}}\sum_{j=-\tilde{n}}^{\tilde{n}}\zeta_j\varphi_{\theta_0,j}\left(I^T_{\varepsilon,j} - 1\right) = O_p(1).$$

Finally, by mean value theorem, the norm of the first term of (5.10) is bounded by

$$CN^{1/2}\left\|\widehat{\theta}-\theta_0\right\|^2\frac{1}{N}\sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]}\frac{I_{x,j}^T}{\left|\Psi_{\theta_0,j}\right|^2}=O_p\left(N^{-1/2}\right), \tag{5.11}$$

by Theorem 2 and proceeding as with the third term of (5.10). This concludes the proof. □

We now introduce the following notation. $\varepsilon^T(t)=h(t)\varepsilon(t)$ and for $v_1<v_2\in[0,\pi]^d$,

$$\mathcal{E}_{1,N}(v_1,v_2)=\left(\frac{1}{N}\sum_{j=[\tilde{n}v_1/\pi]}^{[\tilde{n}v_2/\pi]}\zeta_j\right)\left(\frac{N^{1/2}}{\sum_{t=1}^n h^2(t)}\sum_{t=1}^n\left(\varepsilon^T(t)^2-1\right)\right) \tag{5.12}$$

$$\mathcal{E}_{2,N}(v_1,v_2)=\frac{1}{N}\sum_{j=[\tilde{n}v_1/\pi]}^{[\tilde{n}v_2/\pi]}\zeta_j\frac{N^{1/2}}{\sum_{t=1}^n h^2(t)}\sum_{t_1\neq t_2=1}^n\varepsilon^T(t_1)\varepsilon^T(t_2)e^{i(t_1-t_2)\cdot\lambda_j}. \tag{5.13}$$

Notice that $\mathcal{E}_{1,N}(v_1,v_2)+\mathcal{E}_{2,N}(v_1,v_2)=N^{-1/2}\sum_{j=[\tilde{n}v_1/\pi]}^{[\tilde{n}v_2/\pi]}\zeta_j\left(I_{\varepsilon,j}^T-1\right)$.

Lemma 8. *Let $v_1<v<v_2\in[0,\pi]^d$ and $\zeta(\lambda)$ as in Lemma 3. Then, assuming $C1-C3$, for some $\beta>0$,*

$$\mathbb{E}\left(\left|\mathcal{E}_{j,N}(v_1,v)\right|^\beta\left|\mathcal{E}_{j,N}(v,v_2)\right|^\beta\right)\le C\prod_{\ell=1}^d\left(v_2[\ell]-v_1[\ell]\right)^2,\ j=1,2. \tag{5.14}$$

Proof. The proof follows proceeding as that of Lemma 6 of Delgado et al. (2005) and observing that by continuity of $\zeta(\lambda)$, $\left|N^{-1}\sum_{p=[\tilde{n}v_1/\pi]}^{[\tilde{n}v_2/\pi]}\zeta_p^q\right|\le C\prod_{\ell=1}^d\left(v_2[\ell]-v_1[\ell]\right)$ for any $q\ge1$. □

Next we will show that the processes $\mathcal{E}_{1,N}(0,\lambda)$ and $\mathcal{E}_{2,N}(0,\lambda)$ are tight. From Bickel and Wichura (1971) it suffices to show the following lemma.

Lemma 9. *Assuming $C1$, we have that*

$$(a)\ \mathbb{E}\prod_{\ell=1}^d\left(\mathcal{E}_{1,N}^{(\ell)}\left(0,\lambda_{1[\ell]}\right)-\mathcal{E}_{1,N}^{(\ell)}\left(0,\lambda_{2[\ell]}\right)\right)^2\le C\prod_{\ell=1}^d\left(\lambda_{2[\ell]}-\lambda_{1[\ell]}\right)^2 \tag{5.15}$$

$$(b)\ \mathbb{E}\prod_{\ell=1}^d\left(\mathcal{E}_{2,N}^{(\ell)}\left(0,\lambda_{1[\ell]}\right)-\mathcal{E}_{2,N}^{(\ell)}\left(0,\lambda_{2[\ell]}\right)\right)^4\le C\prod_{\ell=1}^d\left(\lambda_{2[\ell]}-\lambda_{1[\ell]}\right)^2 \tag{5.16}$$

for all $\lambda_{1[\ell]}<\lambda_{2[\ell]}\in[0,\pi]$, $\ell=1,...,d$, and where, say,

$$\mathcal{E}_{1,N}^{(\ell)}\left(\lambda_{1[\ell]},\lambda_{2[\ell]}\right)=\left(\frac{1}{n[\ell]}\sum_{j[\ell]=\left[\tilde{n}\lambda_{1[\ell]}/\pi\right]}^{\left[\tilde{n}\lambda_{2[\ell]}/\pi\right]}\zeta_j\right)\left(\frac{N^{1/2}}{\sum_{t=1}^n h^2(t)}\sum_{t=1}^n\left(\varepsilon^T(t)^2-1\right)\right)$$

$$\mathcal{E}_{2,N}^{(\ell)}\left(\lambda_{1[\ell]},\lambda_{2[\ell]}\right)=\frac{1}{n[\ell]}\sum_{j[\ell]=\left[\tilde{n}\lambda_{1[\ell]}/\pi\right]}^{\left[\tilde{n}\lambda_{2[\ell]}/\pi\right]}\zeta_j\frac{N^{1/2}}{\sum_{t=1}^n h^2(t)}\sum_{t_1\neq t_2=1}^n\varepsilon^T(t_1)\varepsilon^T(t_2)e^{i(t_1-t_2)\cdot\lambda_j}.$$

Proof. The proof follows after observing that $\mathcal{E}_{\ell,N}^{(\ell)}\left(0,\lambda_{1[\ell]}\right)-\mathcal{E}_{\ell,N}^{(\ell)}\left(0,\lambda_{2[\ell]}\right)=\mathcal{E}_{\ell,N}^{(\ell)}\left(\lambda_{1[\ell]},\lambda_{2[\ell]}\right)$ for $\ell=1,2$ and then by Lemma 8. □

Lemma 10. *Under $C1-C3$ and $C5-C8$,*

$$\widehat{\sigma}_\varepsilon^2=\frac{1}{N}\sum_{t=1}^n\widehat{\varepsilon}^2(t)\quad\text{and}\quad\widehat{\mu}_{4,\varepsilon}=\frac{1}{N}\sum_{t=1}^n\widehat{\varepsilon}^4(t)$$

are consistent estimators of σ_ε^2 and $\mu_{4,\varepsilon}$, respectively.

Proof. See Theorem 4 of Robinson and Vidal-Sanz (2006). □

In what follows, we shall abbreviate $w_{x^*}^T(\lambda)/\Psi_{\widehat{\theta}}(\lambda)$ and $w_{\varepsilon^*}^T(\lambda)$ by $u^*(\lambda)$ and $v^*(\lambda)$ respectively for all $\lambda \in \Pi^d$.

Lemma 11. *Consider $j \in \widetilde{\mathbb{Z}}^d$. Then, for $\zeta_{x^*,j}^T$ and $\xi_{x^*,j}^T$ equal to u_j^* or v_j^*,*

$$(a)\ \mathbb{E}^*\left(\zeta_{x^*,j}^T\overline{\xi}_{x^*,j}^T\right) - \widehat{\sigma}_\varepsilon^2 = O_p\left(\vec{n}^{-2}\right); \quad (b)\ \mathbb{E}^*\left(\zeta_{x^*,j}^T\xi_{x^*,j}^T\right) = 0.$$

Proof. We shall handle the case when both $\zeta_{x^*,j}^T$ and $\xi_{x^*,j}^T$ are u_j^*, being the other cases identically handled. We begin with part (a). By definition and using (2.8), it is easy to show that

$$w_{x^*,j}^T = \frac{1}{6^{d/2}}\sum_{k=-\tilde{n}}^{\tilde{n}} \Psi_{\widehat{\theta},k} w_{\varepsilon^*,k} \prod\nolimits_{\ell=1}^d \overline{\mathcal{I}}_\ell(j,k),$$

where $\overline{\mathcal{I}}_\ell(j,k) = 2\mathcal{I}(j[\ell] = k[\ell]) - \mathcal{I}(j[\ell]-1 = k[\ell]) - \mathcal{I}(j[\ell]+1 = k[\ell])$. So, because $\mathbb{E}^*(w_{\varepsilon^*,j}\overline{w}_{\varepsilon^*,k}) = \widehat{\sigma}_\varepsilon^2\mathcal{I}(j=k)$, the left side of the equality in (a) is

$$\widehat{\sigma}_\varepsilon^2\left(\sum_{k=-\tilde{n}}^{\tilde{n}} \left|\Psi_{\widehat{\theta},j}\right|^{-2}\left|\Psi_{\widehat{\theta},k}\right|^2\left\{\prod\nolimits_{\ell=1}^d \overline{\mathcal{I}}_\ell(j,k)\right\}^2 - 1\right)$$

From here the conclusion is standard because $|\Psi_\theta(\lambda)|^2$ is twice differentiable uniformly in $\theta \in \Theta$ for all $\lambda \in \Pi^d$ and that $\widehat{\theta} - \theta_0 = o_p(1)$. Next we show part (b). That follows immediately because, say, $\mathbb{E}^*(w_{\varepsilon^*,k}w_{\varepsilon^*,k}) = 0$. □

Lemma 12. *Let $k \prec j \in \widetilde{\mathbb{Z}}^d$. Then,under $C1-C4$,*

$$\begin{aligned} (a)\quad \mathbb{E}^*\left(\zeta_{x^*,j}^T\overline{\xi}_{x^*,k}^T\right) &= \widehat{\sigma}_\varepsilon^2\left(1+O_p\left(\vec{n}^{-1}\right)\right)\mathcal{I}(|j[\ell]-k[\ell]| = 2, \ell = 1,...,d) \\ (b)\quad \mathbb{E}^*\left(\zeta_{x^*,j}^T\xi_{x^*,k}^T\right) &= 0. \end{aligned}$$

Proof. Again, we shall handle the case when $\zeta_{x^*,j}^T$ and $\xi_{x^*,j}^T$ are u_j^*. We shall examine part (a) only, being part (b) identical. Proceeding as with the proof of part (a) of the previous lemma, the left side of the equality in (a) is

$$\widehat{\sigma}_\varepsilon^2\sum_{p=-\tilde{n}}^{\tilde{n}} \left|\Psi_{\widehat{\theta},j}\right|^{-2}\left|\Psi_{\widehat{\theta},p}\right|^2\left\{\prod\nolimits_{\ell=1}^d \overline{\mathcal{I}}_\ell(j,p)\,\overline{\mathcal{I}}_\ell(k,p)\right\}.$$

From here, we see that the last expression is zero except when $|j[\ell]-k[\ell]| = 2$, for all $\ell = 1,...,d$, in which case is $\left|\Psi_{\widehat{\theta},j\pm1}\right|^2/\left|\Psi_{\widehat{\theta},j}\right|^2 = 1 + O_p\left(\vec{n}^{-1}\right)$ as $|\Psi_\theta(\lambda)|^2$ is twice differentiable uniformly in $\theta \in \Theta$ for all $\lambda \in \Pi^d$ and $\widehat{\theta} - \theta_0 = o_p(1)$. □

Lemma 13. *Let $\zeta(\lambda)$ be as in Lemma 3. Under $C1-C4$, we have that for all $r \le s \in \widetilde{\mathbb{Z}}^d$*

$$\mathbb{E}^*\left|\sum_{j=r}^{s}\zeta_j v_j^*\left(u_j^* - v_j^*\right)\right|^2 = O_p\left(\frac{1}{\vec{n}}\prod_{\ell=1}^d |s[\ell]-r[\ell]|_+\right). \tag{5.17}$$

Proof. Denote $\varrho_j^* = u_j^* - v_j^*$. By standard arguments, the left side of (5.17) is

$$\begin{aligned} &\sum_{j=r}^{s}\zeta_j^2\mathbb{E}^*\left\{v_j^*\overline{v}_j^*\varrho_j^*\overline{\varrho}_j^*\right\} + \sum_{j\neq k=r}^{s}\zeta_j\zeta_k\mathbb{E}\left\{v_j^*\overline{v}_k^*\varrho_j^*\overline{\varrho}_k^*\right\} \\ =\ &\sum_{j=r}^{s}\zeta_j^2\left\{a_{j1}^* + a_{j2}^*\right\} + \sum_{j\neq k=r}^{s}\zeta_j\zeta_k\left\{b_{jk,1}^* + b_{jk,2}^*\right\}, \end{aligned}$$

where, by Lemmas 11 and 12 part (*b*),

$$\begin{aligned}
a^*_{j1} &= \mathbb{E}^*\left(v^*_j\overline{v}^*_j\right)\mathbb{E}^*\left(\varrho^*_j\overline{\varrho}^*_j\right)+\left|\mathbb{E}^*\left(v^*_j\overline{\varrho}^*_j\right)\right|^2;\\
a^*_{j2} &= cum^*\left(v^*_j,\overline{v}^*_j,\overline{u}^*_j,u^*_j\right)+cum^*\left(v^*_j,\overline{v}^*_j,\overline{v}^*_j,v^*_j\right)\\
&\quad -cum^*\left(v^*_j,\overline{v}^*_j,\overline{u}^*_j,v^*_j\right)-cum^*\left(v^*_j,\overline{v}^*_j,u^*_j,\overline{v}^*_j\right)\\
b^*_{jk,1} &= \mathbb{E}^*\left(v^*_j\overline{v}^*_k\right)\mathbb{E}^*\left(\varrho^*_j\overline{\varrho}^*_k\right)+\mathbb{E}^*\left(v^*_j\varrho^*_j\right)\mathbb{E}^*\left(\overline{v}^*_k\overline{\varrho}^*_k\right)\\
b^*_{jk,2} &= cum^*\left(v^*_j,\overline{v}^*_k,\overline{u}^*_j,u^*_k\right)+cum^*\left(v^*_j,\overline{v}^*_k,\overline{v}^*_j,v^*_k\right)\\
&\quad -cum^*\left(v^*_j,\overline{v}^*_k,\overline{u}^*_j,v^*_k\right)-cum^*\left(v^*_j,\overline{v}^*_k,u^*_j,\overline{v}^*_k\right).
\end{aligned}$$

After observing that $\mathbb{E}^*\left(v^*_j\overline{u}^*_j\right)=\widehat{\sigma}^2_\varepsilon+O_p\left(\vec{n}^{-2}\right)$ and $\mathbb{E}^*\left(v^*_j\overline{\varrho}^*_j\right)=\mathbb{E}^*\left(v^*_j\overline{u}^*_j\right)-\mathbb{E}^*\left(v^*_j\overline{v}^*_j\right)$, we have that Lemma 11 implies that $a^*_{j1}=O_p\left(\vec{n}^{-2}\right)$, whereas Lemmas 11 and 12 imply that $b^*_{jk,1}=O_p\left(\vec{n}^{-2}+\vec{n}^{-1}\mathcal{I}\left(\left|j\left[\ell\right]-k\left[\ell\right]\right|=2,\ell=1,...,d\right)\right)$. From here it is immediate to conclude that the contribution due to a^*_{j1} and $b^*_{jk,1}$ into the left of (5.17) is its right side.

Finally we examine a^*_{j2} and $b^*_{jk,2}$. By definition of, for example, $w^T_{x^*}(\lambda)$ and that $cum^*\left(\varepsilon^*(t_1),...,\varepsilon^*(t_4)\right)=\widehat{\kappa}_\varepsilon\mathcal{I}\left(t_1=...=t_4\right)$, it is obvious that $b^*_{jk,2}$ is $O_p\left(n^{-1}\mathcal{I}\left(\left|j\left[\ell\right]-k\left[\ell\right]\right|=2,\ell=1,...,d\right)\right)$. Notice that $\widehat{\kappa}_\varepsilon=\widehat{\mu}_{4,\varepsilon}-3\widehat{\sigma}^4_\varepsilon=O_p(1)$ by Lemma 10. □

Lemma 14. *Let $\zeta(\lambda)$ be a function as in Lemma 3. Then, under C1 − C4,*

$$\mathbb{E}^*\sup_{\lambda\in[0,\pi]^d}\left|\sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]}\zeta_j\left\{\frac{I^T_{x^*,j}}{\left|\Psi_{\widehat{\theta},j}\right|^2}-I^T_{\varepsilon^*,j}\right\}\right|=o_p\left(N^{1/2}\right).$$

Proof. We shall consider the proof in the positive quadrant $\sum_{j=1}^{[\tilde{n}\lambda/\pi]}$, being the proof for the remaining $2^{d-1}-1$ quadrants similarly handled. By the Cauchy-Schwarz and triangle inequalities, it suffices to show that

$$\mathbb{E}^*\sup_s\left|\sum_{j=1}^s\zeta_j\left\{\frac{I^T_{x^*,j}}{\left|\Psi_{\widehat{\theta},j}\right|^2}-I^T_{\varepsilon^*,j}\right\}\right|\leq\mathbb{E}^*\sup_s\sum_{j=1}^s\left|\zeta_j\right|\left|\varrho^*_j\right|^2+2\mathbb{E}^*\sup_s\left|\sum_{j=1}^s\zeta_jv^*_j\overline{\varrho}^*_j\right| \tag{5.18}$$

is $o_{p^*}\left(N^{1/2}\right)$, where we abbreviate "$\sup_{s=1,...,\tilde{n}}$" by "$\sup_s$" and $\varrho^*_j=u^*_j-v^*_j$.

The first term on the right of (5.18) is bounded by

$$C\sum_{j=1}^{\tilde{n}}\left\{\left(\mathbb{E}^*\left|u^*_j\right|^2-\widehat{\sigma}^2_\varepsilon\right)-\left(\mathbb{E}^*\left(u^*_j\overline{v}^*_j\right)-\widehat{\sigma}^2_\varepsilon\right)-\left(\mathbb{E}^*\left(\overline{u}^*_jv^*_j\right)-\widehat{\sigma}^2_\varepsilon\right)+\left(\mathbb{E}^*\left|v^*_j\right|^2-\widehat{\sigma}^2_\varepsilon\right)\right\}=o_p\left(N^{1/2}\right),$$

because $\left|\zeta_j\right|\leq C$, $d<4$ and by Lemma 11, for instance

$$\left|\mathbb{E}^*\left(u^*_j\left\{\begin{array}{c}\overline{v}^*_j\\\overline{u}^*_j\end{array}\right.\right)-\widehat{\sigma}^2_\varepsilon\right|\leq\widehat{\sigma}^{-2}_\varepsilon\left|\Psi_{\widehat{\theta},j}\right|^{-2}\left(\left|\mathbb{E}^*\left(w^T_{x^*,j}\left\{\begin{array}{c}\overline{w}^T_{\varepsilon^*,j}\\\overline{w}^T_{x^*,j}\end{array}\right.\right)-\widehat{\sigma}^2_\varepsilon\left\{\begin{array}{c}\Psi_{\widehat{\theta},j}\\\left|\Psi_{\widehat{\theta},j}\right|^2\end{array}\right.\right|\right)=O_p\left(\frac{1}{\vec{n}^2}\right).$$

Next, we examine the second term of (5.18). With same notation as in Lemma 4, the square of the second term on the right of (5.18) is bounded by

$$\mathbb{E}^*\max_s\left|\left\{\sum_{j=1}^s-\sum_{j=1}^{q(s)\left[\tilde{n}^{1-\varsigma}\right]}\right\}\zeta_jv^*_j\overline{\varrho}^*_j\right|^2+\mathbb{E}^*\max_s\left|\sum_{j=1}^{q(s)\left[\tilde{n}^{1-\varsigma}\right]}\zeta_jv^*_j\overline{\varrho}^*_j\right|^2. \tag{5.19}$$

From the definition of $q(s)$ and $\left(\sup_p |c_p|\right)^2 = \sup_p |c_p|^2 \leq \sum_p |c_p|^2$, the second term of (5.19) is bounded by $\sum_{q=1}^{[\tilde{n}^{\varsigma}]-1} \mathbb{E}^* \left| \sum_{j=1}^{q[\tilde{n}^{1-\varsigma}]} \zeta_j v_j^* \overline{\varrho}_j^* \right|^2 = O_p\left(N^{1+\varsigma}\vec{n}^{-1} \log^2 \vec{n}\right) = o_p(N)$ by Lemma 13 and because $\varsigma < 1/d$.

To complete the proof we need to show that the first term in (5.19) is $o_p(N)$. To that end, we note that it is bounded by

$$\mathbb{E}^* \max_{q=1,\ldots,[\tilde{n}^{\varsigma}]-1} \max_{s=1+q[\tilde{n}^{1-\varsigma}],\ldots,(q+1)[\tilde{n}^{1-\varsigma}]} \left| \sum_{j=1+q[\tilde{n}^{1-\varsigma}]}^{s} \zeta_j v_j^* \overline{\varrho}_j^* \right|^2$$

which is $O_p(N^{\varsigma}) \mathbb{E}^* \max_{s=1,\ldots,[\tilde{n}^{1-\varsigma}]} \left| \sum_{j=1}^{s} \zeta_j v_j^* \overline{\varrho}_j^* \right|^2$. So, we have that the square of the second term on the right of (5.18) is

$$o_p(N) + O_p(N^{\varsigma}) \mathbb{E}^* \max_{s=1,\ldots,[\tilde{n}^{1-\varsigma}]} \left| \sum_{j=1}^{s} \zeta_j v_j^* \overline{\varrho}_j^* \right|^2 .$$

Now proceeding as in Lemma 4, the square of the second term on the right of (5.18) is

$$o_p(N) + O_p\left(N^{\varsigma \sum_{p=0}^{\iota-1}(1-\varsigma)^p}\right) \sum_{s=1}^{\tilde{n}^{(1-\varsigma)^{\iota}}} \mathbb{E}^* \left| \sum_{j=1}^{s} \zeta_j v_j^* \overline{\varrho}_j^* \right|^2 = o_p(N)$$

after choosing ι large enough because $\varsigma < 1/d$. This completes the proof. □

Lemma 15. *Let $\zeta(\lambda;\vartheta)$ be is as in Lemma 5. Assuming $C1-C4$,*

$$\frac{1}{N} \sup_{\vartheta \in \Theta \times \mathbb{R}^+} \left| \sum_{j=-\tilde{n}}^{\tilde{n}} \zeta_j(\vartheta) \left(\frac{I_{x^*,j}^T}{\left|\Psi_{\widehat{\theta},j}\right|^2} - \widehat{\sigma}_{\varepsilon}^2 \right) \right| = o_{p^*}(1) . \tag{5.20}$$

Proof. By the triangle inequality, the left side of (5.20) is bounded by

$$\frac{C}{N} \sup_{\vartheta \in \Theta \times \mathbb{R}^+} \left| \sum_{j=-\tilde{n}}^{\tilde{n}} \zeta_j(\vartheta) \left(\frac{I_{x^*,j}^T}{\left|\Psi_{\widehat{\theta},j}\right|^2} - I_{\varepsilon^*,j}^T \right) \right| + \frac{C}{N} \sup_{\vartheta \in \Theta \times \mathbb{R}^+} \left| \sum_{j=-\tilde{n}}^{\tilde{n}} \zeta_j(\vartheta) \left(I_{\varepsilon^*,j}^T - \widehat{\sigma}_{\varepsilon}^2 \right) \right| . \tag{5.21}$$

Now, because by assumption $|\zeta(\lambda;\vartheta)| < C$, the first term of (5.21) is bounded by

$$\frac{C}{N} \sum_{j=-\tilde{n}}^{\tilde{n}} \left| \frac{I_{x^*,j}^T}{\left|\Psi_{\widehat{\theta},j}\right|^2} - I_{\varepsilon^*,j}^T \right| \leq \frac{C}{N} \sum_{j=-\tilde{n}}^{\tilde{n}} \left| u_j^* - v_j^* \right|^2 + \frac{C}{N} \sum_{j=-\tilde{n}}^{\tilde{n}} \left| v_j^* \left(\overline{u}_j^* - \overline{v}_j^* \right) \right| = o_{p^*}(1)$$

by Markov's inequality because by the Cauchy-Schwarz inequality $\mathbb{E}^* \left| v_j^* \left(\overline{u}_j^* - \overline{v}_j^* \right) \right|^2 \leq \mathbb{E}^* \left| v_j^* \right|^2 \mathbb{E}^* \left| \overline{u}_j^* - \overline{v}_j^* \right|^2$ and then proceeding as in Lemma 13. Next, the second term of (5.21) is $o_{p^*}(1)$. First, the finite dimensional distributions of $S^*(\vartheta) = N^{-1} \sum_{j=-\tilde{n}}^{\tilde{n}} \zeta_j(\vartheta) \left(I_{\varepsilon^*,j}^T - \widehat{\sigma}_{\varepsilon}^2 \right)$ converge to zero in probability. Indeed, the second bootstrap moment is

$$\frac{1}{N^2} \sum_{j,k=-\tilde{n}}^{\tilde{n}} \zeta_j(\vartheta) \zeta_k(\vartheta) \mathbb{E}^* \left\{ \left(I_{\varepsilon^*,j}^T - \widehat{\sigma}_{\varepsilon}^2 \right) \left(I_{\varepsilon^*,k}^T - \widehat{\sigma}_{\varepsilon}^2 \right) \right\} = O_p\left(\frac{1}{N}\right)$$

by standard algebra after observing that $\mathbb{E}^*(w_{\varepsilon^*,j} \overline{w}_{\varepsilon^*,k}) = \widehat{\sigma}_{\varepsilon}^2 \mathcal{I}(j=k)$, $\mathbb{E}^*(w_{\varepsilon^*,j} w_{\varepsilon^*,k}) = 0$ and $cum^*(w_{\varepsilon^*,j}, \overline{w}_{\varepsilon^*,j}, w_{\varepsilon^*,k}, \overline{w}_{\varepsilon^*,k}) = O\left(N^{-1}\right) \widehat{\kappa}_{\varepsilon} \mathcal{I}(j=k)$. To finish, we need to show the tightness of the process $S^*(\vartheta)$. But this is immediate because proceeding as with the last displayed equality

$$\mathbb{E}^* \left|S^*(\vartheta_2) - S^*(\vartheta_1)\right|^2 = |\vartheta_2 - \vartheta_1|^2 O_p(1)$$

because continuous differentiability of $\zeta(\lambda;\vartheta)$ for all $\lambda \in \Pi^d$ implies that $\left|\zeta_j(\vartheta_2) - \zeta_j(\vartheta_1)\right| \leq C|\vartheta_2 - \vartheta_1|$. □

Lemma 16. *Assume $C1 - C3$ and $C5 - C8$. Then, $\widehat{\vartheta}^* - \widehat{\vartheta} = o_{p^*}(1)$.*

Proof. The proof follows very easily using Lemma 15. Indeed, (4.5) is

$$\frac{1}{N}\sum_{j=-\tilde{n}}^{\tilde{n}} \frac{f_{\widehat{\vartheta},j}}{f_{\vartheta,j}}\left(\frac{I^T_{x^*,j}}{(2\pi)^d f_{\widehat{\vartheta},j}} - 1\right) + \frac{1}{N}\left\{\sum_{j=-\tilde{n}}^{\tilde{n}} \frac{f_{\widehat{\vartheta},j}}{f_{\vartheta,j}} - \log\frac{f_{\widehat{\vartheta},j}}{f_{\vartheta,j}} + \log f_{\widehat{\vartheta},j}\right\}. \tag{5.22}$$

Now, the difference between the second term of (5.22) and

$$\int_{\pi}^{\pi}\left\{\frac{f_{\widehat{\vartheta}}(\lambda)}{f_{\vartheta}(\lambda)} - \log\left(\frac{f_{\widehat{\vartheta}}(\lambda)}{f_{\vartheta}(\lambda)}\right)\right\}d\lambda + \int_{\pi}^{\pi}\log f_{\widehat{\vartheta}}(\lambda)\,d\lambda$$

converges to zero in probability using Brillinger (1981, p.15) and that uniformly in λ, $\left|f_{\widehat{\vartheta}}(\lambda) - f_{\vartheta_0}(\lambda)\right| = o_p(1)$. Moreover, the last displayed expression is greater than or equal to $\frac{(2\pi)^d}{2} + \int_{\pi}^{\pi}\log f_{\widehat{\vartheta}}(\lambda)\,d\lambda$ with equality when $f_{\widehat{\vartheta}}(\lambda) = f_{\vartheta}(\lambda)$ which is the case only if $\vartheta = \widehat{\vartheta}$ by $C7$. On the other hand, the first term of (5.22) converges to zero uniformly in ϑ by Lemma 15 because $f^{-1}_{\vartheta,j}f_{\widehat{\vartheta},j}$ satisfies the same conditions as $\zeta(\lambda;\vartheta)$ there by $C6$. From here the conclusion of the lemma is standard proceeding as in Theorem 1 of Hannan (1973), so we omit its details. □

Lemma 17. *Assume $C1 - C8$. Under H_0, uniform in $\lambda \in [0,\pi]^d$*

$$\frac{1}{N^{1/2}}\sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]}\zeta_j\left(\frac{I^T_{x^*,j}}{\left|\Psi_{\widehat{\theta}^*,j}\right|^2} - I^T_{\varepsilon^*,j}\right) = -\left(\frac{1}{N}\sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]}\zeta_j\varphi'_{\widehat{\theta},j}\right)N^{1/2}\left(\widehat{\theta}^* - \widehat{\theta}\right) + o_{p^*}(1), \tag{5.23}$$

where $\zeta(\lambda)$ is as in Lemma 3.

Proof. The difference between the left side of (5.23) and the first term on its right side is

$$\begin{aligned}&\frac{1}{N^{1/2}}\sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]}\zeta_j\frac{I^T_{x^*,j}}{\left|\Psi_{\widehat{\theta},j}\right|^2}\left[\frac{\left|\Psi_{\widehat{\theta},j}\right|^2}{\left|\Psi_{\widehat{\theta}^*,j}\right|^2} - 1 + \varphi'_{\widehat{\theta},j}\left(\widehat{\theta}^* - \widehat{\theta}\right)\right]\\ &+\frac{1}{N^{1/2}}\sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]}\zeta_j\left(\frac{I^T_{x^*,j}}{\left|\Psi_{\widehat{\theta},j}\right|^2} - I^T_{\varepsilon^*,j}\right) - \frac{1}{N^{1/2}}\sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]}\zeta_j\varphi'_{\widehat{\theta},j}\frac{I^T_{x^*,j}}{\left|\Psi_{\widehat{\theta},j}\right|^2}\left(\widehat{\theta}^* - \widehat{\theta}\right).\end{aligned} \tag{5.24}$$

First, because each component of the vector $\zeta(\lambda)\varphi_{\widehat{\theta}}(\lambda)$ satisfies the same conditions of $\zeta(\lambda)$ in Lemma 14, Markov's inequality implies that the second term of (5.24) is $o_{p^*}(1)$, whereas the third term is $N^{-1}\sum_{j=1}^{[\tilde{n}\lambda/\pi]}\zeta_j\varphi'_{\widehat{\theta},j}N^{1/2}\left(\widehat{\theta}^* - \widehat{\theta}\right) + o_{p^*}(1)$ by Lemma 14 and because proceeding as in the proof of Theorem 4

$$\frac{1}{N^{1/2}}\sum_{j=-\tilde{n}}^{\tilde{n}}\zeta_j\varphi_{\theta,j}\left(I^T_{\varepsilon^*,j} - \widehat{\sigma}^2_{\varepsilon}\right) = O_{p^*}(1)$$

Finally, by mean value theorem, the norm of the first term of (5.24) is bounded by

$$CN^{1/2}\left\|\widehat{\theta}^{*}-\widehat{\theta}\right\|^{2}\frac{1}{N}\sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]}\frac{I_{x^{*},j}^{T}}{\left|\Psi_{\widehat{\theta}^{*},j}\right|^{2}}=O_{p^{*}}\left(N^{-1/2}\right),$$

by Theorem 4 and proceeding as with the second term of (5.24). This concludes the proof. □

We now introduce the following notation. For $v_1, v_2 \in [0,\pi]^d$, denote $\mathcal{E}_{1,N}^{*}(v_1,v_2)$ and $\mathcal{E}_{2,N}^{*}(v_1,v_2)$ as in (5.12) and (5.13) but with $\varepsilon^{T}(t)$ there replaced by $\varepsilon^{*T}(t) = h(t)\,\varepsilon^{*}(t)$ and also let $H_N = H_N(v_1,v_2)$ a sequence bounded in probability.

Lemma 18. *Let $v_1 < v < v_2 \in [0,\pi]^d$. Then, assuming $C1-C3$, and for some $\beta > 0$, with $\zeta(\lambda)$ as in Lemma 3,*

$$\mathbb{E}^{*}\left(\left|\mathcal{E}_{j,N}^{*}(v_1,v)\right|^{\beta}\left|\mathcal{E}_{j,N}^{*}(v,v_2)\right|^{\beta}\right)\le H_N(v_1,v_2)\prod_{\ell=1}^{d}\left(v_2[\ell]-v_1[\ell]\right)^2,\ j=1,2.$$

Proof. The proof proceeds as in Lemma 8 but instead of using Delgado et al.' (2005) Lemma 5 we use Lemma 7.3 of Hidalgo and Kreiss (2006). □

Next we will show that the processes $\mathcal{E}_{1,N}^{*}(0,\lambda)$ and $\mathcal{E}_{2,N}^{*}(0,\lambda)$ are tight. To that end, it suffices to show the following lemma.

Lemma 19. *Assuming $C1$ we have that*

(5.25)

$$(a)\ \ \mathbb{E}^{*}\prod_{\ell=1}^{d}\left(\mathcal{E}_{1,N}^{(\ell)*}\left(0,\lambda_{1[\ell]}\right)-\mathcal{E}_{1,N}^{(\ell)*}\left(0,\lambda_{2[\ell]}\right)\right)^{2}=H_N(\lambda_1,\lambda_2)\prod_{\ell=1}^{d}\left(\lambda_{2[\ell]}-\lambda_{1[\ell]}\right)^{2}$$

(5.26)

$$(b)\ \ \mathbb{E}^{*}\prod_{\ell=1}^{d}\left(\mathcal{E}_{2,N}^{(\ell)*}\left(0,\lambda_{1[\ell]}\right)-\mathcal{E}_{2,N}^{(\ell)*}\left(0,\lambda_{2[\ell]}\right)\right)^{4}=H_N(\lambda_1,\lambda_2)\prod_{\ell=1}^{d}\left(\lambda_{2[\ell]}-\lambda_{1[\ell]}\right)^{2}$$

for all $\lambda_{1[\ell]} < \lambda_{2[\ell]} \in [0,\pi]$, $\ell = 1,...,d$, and where

$$\mathcal{E}_{1,N}^{(\ell)*}\left(\lambda_{1[\ell]},\lambda_{2[\ell]}\right)=\left(\frac{1}{n[\ell]}\sum_{j[\ell]=\left[\tilde{n}\lambda_{1[\ell]}/\pi\right]}^{\left[\tilde{n}\lambda_{2[\ell]}/\pi\right]}\zeta_j\right)\left(\frac{N^{1/2}}{\sum_{t=1}^{n}h^2(t)}\sum_{t=1}^{n}\left(\varepsilon^{*T}(t)^2-1\right)\right)$$

$$\mathcal{E}_{2,N}^{(\ell)*}\left(\lambda_{1[\ell]},\lambda_{2[\ell]}\right)=\frac{1}{n[\ell]}\left(\sum_{j[\ell]=\left[\tilde{n}\lambda_{1[\ell]}/\pi\right]}^{\left[\tilde{n}\lambda_{2[\ell]}/\pi\right]}\zeta_j\frac{N^{1/2}}{\sum_{t=1}^{n}h^2(t)}\sum_{t_1\ne t_2=1}^{n}\varepsilon^{*T}(t_1)\,\varepsilon^{*T}(t_2)\,e^{i(t_1-t_2)\cdot\lambda_j}\right).$$

Proof. The proof follows after observing that $\mathcal{E}_{\ell,N}^{(\ell)*}\left(0,\lambda_{1[\ell]}\right)-\mathcal{E}_{\ell,N}^{(\ell)*}\left(0,\lambda_{2[\ell]}\right)=\mathcal{E}_{\ell,N}^{(\ell)*}\left(\lambda_{1[\ell]},\lambda_{2[\ell]}\right)$ for $\ell = 1,2$ and then by Lemma 18. □

6. PROOFS

6.1. Proof of Proposition 1.

We shall be a bit more general. In particular, for a vector function $\zeta(\lambda)$ as in Lemma 3, we will show that

$$\boldsymbol{S}_N(\mu)=\frac{1}{N^{1/2}}\sum_{j=-[\tilde{n}\mu/\pi]}^{[\tilde{n}\mu/\pi]}\zeta_j\left(I_{\varepsilon,j}^{T}-1\right)\Rightarrow\boldsymbol{B}_{\zeta}(\mu),\qquad\mu\in[0,\pi]^d$$

where for $\mu \le v \in [0,\pi]^d$, $Cov\left(\boldsymbol{B}_{\zeta}(\mu),\boldsymbol{B}_{\zeta}(v)\right)=\left(2+\kappa_{\varepsilon}\left(\frac{35}{18}\right)^{d}\right)\int_{-\mu}^{\mu}\zeta(\lambda)\,\zeta'(\lambda)\,d\lambda$, as $n\left(\sum_{t=1}^{n}h_t^4\right)/\left(\sum_{t=1}^{n}h_t^2\right)^2=\frac{35}{18}$.

To that end, it suffices to show that (a) for all $\mu \in [0,\pi]^d$,
(6.1)
$$\frac{1}{N^{1/2}}\sum_{j=-[\tilde{n}\mu/\pi]}^{[\tilde{n}\mu/\pi]}\zeta_j\left(I^T_{\varepsilon,j}-1\right)\overset{d}{\to}\mathcal{N}\left(0,\left(2+\kappa_\varepsilon\left(\frac{35}{18}\right)^d\right)\int_{-\mu}^{\mu}\zeta(\lambda)\zeta'(\lambda)\,d\lambda\right).$$

(b) For $\mu \le v$, the covariance structure is such that
$$\frac{1}{N}E\left\{\sum_{j=-[\tilde{n}\mu/\pi]}^{[\tilde{n}\mu/\pi]}\zeta_j\left(I^T_{\varepsilon,j}-1\right)\sum_{j=-[\tilde{n}v/\pi]}^{[\tilde{n}v/\pi]}\zeta_j\left(I^T_{\varepsilon,j}-1\right)\right\}\to\left(2+\kappa_\varepsilon\left(\frac{35}{18}\right)^d\right)\int_{-\mu}^{\mu}\zeta(\lambda)\zeta'(\lambda)\,d\lambda.$$

(c) the process $\left\{\boldsymbol{S}_N(\mu):\mu\in[0,\pi]^d\right\}$ is tight.

We begin with (a). Its proof follows directly by that in Robinson and Vidal-Sanz (2006) and observing that because $\zeta(\lambda)$ is continuously differentiable, then by Brillinger (1981, p.15), $N^{-1}\sum_{j=-[\tilde{n}v/\pi]}^{[\tilde{n}v/\pi]}\zeta_j\zeta_j'-(2\pi)^{-d}\int_{-v}^{v}\zeta(\lambda)\zeta'(\lambda)\,d\lambda=O_p\left(\vec{n}^{-1}\right)$, and thus it is omitted.

Part (b) follows after observing that $\mathbb{E}I^T_{\varepsilon,j}=1$ by $C1$ and $\mathbb{E}\left(I^T_{\varepsilon,j}I^T_{\varepsilon,k}\right)$ is
$$\begin{aligned}&\frac{1}{\left(\sum_{t=1}^{n}h^2(t)\right)^2}\sum_{t,s,r,u=1}^{n}\mathbb{E}\left\{\varepsilon^T(t)\,\varepsilon^T(s)\,\varepsilon^T(r)\,\varepsilon^T(s)\right\}e^{-i(t-s)\cdot\lambda_j+i(r-u)\cdot\lambda_k}\\ =\;&2\mathcal{I}(j-k=0,n)+\mathcal{I}(j+k=0,n)+\frac{1}{N}\left(2+\kappa_\varepsilon\left(\frac{35}{18}\right)^d\right)\end{aligned}$$
using that, say $\sum_{p[\ell]=1}^{n[\ell]}h_\ell(t[\ell])\,e^{-ip[\ell]\left(\lambda_{j[\ell]\pm k[\ell]}\right)}=n[\ell]\,\mathcal{I}(j[\ell]\pm k[\ell]=0,n[\ell])$ and that by Brillinger (1981, p.15) we have that $N^{-1}\sum_{t=1}^{\tilde{n}}h^p(t)\to 2^d\int_{[0,1]^d}h^p(u)\,du$ for all $p\ge 0$. Finally, part (c) follows by Lemma 9.

6.2. **Proof of Theorem 3.**

Part (a). The proof is identical to that of Theorem 5, but instead of using Lemmas 14 and 17 we employ respectively Lemmas 4 and 7. Next the proof of part (b) is identical to that of Theorem 5, but instead of Proposition 1 we employ Theorem 4 and instead of Lemmas 18 and 19 we employ Lemmas 8 and 9. □

6.3. **Proof of Theorem 4.**

First, by standard algebra $\widehat{\vartheta}^*-\widehat{\vartheta}=-\overline{Q}^{*-1}_{\widetilde{\vartheta},N}q^*_{\widehat{\vartheta},N}$, where $\widetilde{\vartheta}^*$ is an intermediate point between $\widehat{\vartheta}$ and $\widehat{\vartheta}^*$,
$$q^*_{\vartheta,N}=\frac{1}{N}\sum_{j=-\tilde{n}}^{\tilde{n}}\phi_{\vartheta,j}\left\{\frac{I^T_{x^*,j}}{\sigma^2_\varepsilon\left|\Psi_{\theta,j}\right|^2}-1\right\}$$
and $\overline{Q}^*_{\vartheta,N}$ is as defined in the proof of Theorem 2 but with $I^T_{x,j}$ replaced by $I^T_{x^*,j}$ there. Now, because $f_{\vartheta,j}$ is twice continuously differentiable by $C6$, and $\widehat{\vartheta}^*-\widehat{\vartheta}=o_{p^*}(1)$ by Lemma 16, we easily conclude that $\left|\overline{Q}^*_{\widetilde{\vartheta}^*,N}-Q_{\widetilde{\vartheta}^*,N}\right|=o_{p^*}(1)$, and that $\left|Q_{\widetilde{\vartheta}^*,N}-Q_{\vartheta_0,N}\right|=o_{p^*}(1)$. On the other hand, by Lemma 14 with $\zeta_j=\phi_{\widehat{\vartheta},j}$ there,
$$q^*_{\widehat{\vartheta},N}=N^{-1/2}\sum_{j=-\tilde{n}}^{\tilde{n}}\phi_{\widehat{\vartheta},j}\left\{I^T_{\varepsilon^*,j}/\widehat{\sigma}^2_\varepsilon-1\right\}+o_{p^*}(1).$$

So, to complete the proof it suffices to show

$$(a)\quad \frac{1}{N}\mathbb{E}^*\left\{\sum_{j=-\tilde{n}}^{\tilde{n}}\phi_{\widehat{\vartheta},j}\left\{\frac{I^T_{\varepsilon^*,j}}{\widehat{\sigma}^2_\varepsilon}-1\right\}\sum_{j=-\tilde{n}}^{\tilde{n}}\phi'_{\widehat{\vartheta},j}\left\{\frac{I^T_{\varepsilon^*,j}}{\widehat{\sigma}^2_\varepsilon}-1\right\}\right\}$$
$$\xrightarrow{P} 2+\kappa_\varepsilon\left(\frac{35}{18}\right)^d\int_{-\mu}^{\mu}\phi_{\vartheta_0}(\lambda)\,\phi'_{\vartheta_0}(\lambda)\,d\lambda,$$

and then (b) the Lindeberg's condition

$$\mathbb{E}^*\left\{\frac{1}{N}\sum_{j=-\tilde{n}}^{\tilde{n}}\left|\frac{I^T_{\varepsilon^*,j}}{\widehat{\sigma}^2_\varepsilon}-1\right|^2\mathcal{I}\left(\left|\frac{I^T_{\varepsilon^*,j}}{\widehat{\sigma}^2_\varepsilon}-1\right|^2>\delta N^{-1}\right)\right\}\xrightarrow{P}0.$$

We begin with (a). By Lemmas 11 and 12, we have that

$$\mathbb{E}^*\left\{\left(\frac{I^T_{\varepsilon^*,j}}{\widehat{\sigma}^2_\varepsilon}-1\right)\left(\frac{I^T_{\varepsilon^*,k}}{\widehat{\sigma}^2_\varepsilon}-1\right)\right\}=\mathcal{I}(j=k\pm 2)+cum^*\left(w_{\varepsilon^*,j},\overline{w}_{\varepsilon^*,j},w_{\varepsilon^*,k},\overline{w}_{\varepsilon^*,k}\right).$$

Now, because $cum^*(\varepsilon^*(t_1),\varepsilon^*(t_2),\varepsilon^*(t_3),\varepsilon^*(t_4))=\widehat{\kappa}_\varepsilon\mathcal{I}(t_1=...=t_4)$, we have that the left side in (a) is

$$\frac{1}{N}\sum_{j=-\tilde{n}}^{\tilde{n}}\phi_{\widehat{\vartheta},j}\phi'_{\widehat{\vartheta},j}\left(2+\widehat{\kappa}_\varepsilon\left(\frac{35}{18}\right)\right).$$

From here we conclude (a) by Lemma 10 and that $\phi_\vartheta(\lambda)$ is continuous by $C6$.

We now show (b). By standard inequalities, the left side is bounded by

$$\frac{1}{N^2}\sum_{j=-\tilde{n}}^{\tilde{n}}\mathbb{E}^*\left|\frac{I^T_{\varepsilon^*,j}}{\widehat{\sigma}^2_\varepsilon}-1\right|^4\leq C\frac{\log^2 N}{N^2}\sum_{j=-\tilde{n}}^{\tilde{n}}\mathbb{E}^*\left|\frac{I^T_{\varepsilon^*,j}}{\widehat{\sigma}^2_\varepsilon}\right|^2$$

by An et al. (1983) because $\{\varepsilon^*(t)\}_{t=1}^n$ is a random sample. Now conclude part (b) since $\mathbb{E}^*\left|I^T_{\varepsilon^*,j}/\widehat{\sigma}^2_\varepsilon\right|^2=O_p(1)$ by Lemma 10. This concludes the proof. □

6.4. **Proof of Theorem 5.**

Part (a). By Lemma 17 with $\zeta(\lambda)=1$ there and the definitions of $G^*_{\theta,N}(\lambda)$ and $G^{*0}_N(\lambda)$, we have that by Theorem 4 uniform in $\lambda\in[0,\pi]^d$,

$$N^{1/2}\left(G^*_{\widehat{\theta}^*,N}(\lambda)-G^{*0}_N(\lambda)\right)=-\left(\frac{1}{N}\sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]}\widetilde{\varphi}'_{\widehat{\theta},N}(j)\right)N^{1/2}\left(\widehat{\theta}^*-\widehat{\theta}\right)+o_{p^*}(1)$$

$$=-\left(\frac{1}{N}\sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]}\widetilde{\varphi}'_{\widehat{\theta},N}(j)\right)\widetilde{\Lambda}^{-1}_{\widehat{\theta},N}\frac{1}{N^{1/2}}\sum_{k=-\tilde{n}}^{\tilde{n}}\widetilde{\varphi}_{\widehat{\theta},N}(k)\frac{I^T_{x^*,k}}{\left|\Psi_{\widehat{\theta},k}\right|^2}+o_{p^*}(1).\qquad(6.2)$$

Now because $\left|G^*_{\widehat{\theta},N}(\pi)-G^{*0}_N(\pi)\right|=o_{p^*}\left(N^{-1/2}\right)$ by Lemma 14, by (6.2), we obtain that uniformly in λ, $\alpha^*_{\widehat{\theta}^*,N}(\lambda)$ is

$$\alpha^{*0}_N(\lambda)+\frac{N^{1/2}\left(G^*_{\widehat{\theta}^*,N}(\lambda)-G^{*0}_N(\lambda)\right)}{G^{*0}_N(\pi)}+G^*_{\widehat{\theta},N}(\lambda)\,N^{1/2}\left(\frac{1}{G^*_{\widehat{\theta}^*,N}(\pi)}-\frac{1}{G^{*0}_N(\pi)}\right)$$

$$=\alpha^{*0}_N(\lambda)-\left(\frac{1}{N}\sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]}\widetilde{\varphi}'_{\widehat{\theta},N}(j)\right)\widetilde{\Lambda}^{-1}_{\widehat{\theta},N}\frac{1}{N^{1/2}}\sum_{k=-\tilde{n}}^{\tilde{n}}\widetilde{\varphi}_{\widehat{\theta},N}(k)\,I^T_{\varepsilon^*,k}+o_{p^*}(1).$$

Now conclude the proof of part (a) by observing that Lemma 10 implies that $\mathbb{E}^* \left| G_N^{*0}(\pi) - \widehat{\sigma}_\epsilon^2 \right| = o_p(1)$ and then by Markov's inequality.

Next part (b). Taking into account part (a), part (b) follows because Theorem 4 guarantees the finite dimensional distributions convergence of α_N^{*0}, whereas its tightness follows by Lemma 19. Tightness of the second term on the right of the last displayed equality follows by Lemmas 18 and 19 because $\varphi_\theta(u)$ is continuously differentiable in θ and u. This concludes the proof. □

References

[1] An, H.-Z., Chen, Z.G. and Hannan, E.J., 1983, The maximum of the periodogram. *Journal of Multivariate Analysis* **13**, 383-400.

[2] Anderson, T.W., 1997, Goodness-of-fit for autoregressive processes. *Journal of Time Series Analysis* **18**, 321-339.

[3] Anderson, T.W. and Walker A.M., 1964, On the asymptotic distribution of the autocorrelations of a sample from a linear stochastic process. *Annals of Mathematical Statistics* **35**, 1296-1303.

[4] Bickel, P.J. and Wichura, M.J., 1971, Convergence criteria for multiparameter stochastic processes and its applications. *Annals of Mathematical Statistics* **42**, 1656-1670.

[5] Billingsley, P., 1968, Convergence of probability measures. Wiley and Sons. New York.

[6] Bloomfield, P., 1973, An exponential model for the spectrum of a scalar time series, *Biometrika* **61**, 217-226.

[7] Box, G.E.P. and Pierce, D.A., 1970, Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of the American Statistical Association* **65**, 1509-1526.

[8] Brillinger, D.R., 1981, Time series, data analysis and theory. Holden-Day. San Francisco.

[9] Chen, H. and Romano J.P.,1999, Bootstrap-assisted goodness-of-fit tests in the frequency domain. *Journal of Time Series Analysis* **20,** 619-654.

[10] Crujeiras, R.M., Fernandez-Casal, R. and Gonzalez-manteiga, W., 2006, Gooodness-of-fit tests for the spatial spectral density. *Preprint*. Universidad de Santiago de Compostela.

[11] Dahlhaus, R., and Künsch, H., 1986, Edge effect and efficient parameter estimation for stationary random fields. *Biometrika* **74**, 877-882.

[12] Delgado, M.A., Hidalgo, J. and Velasco, C., 2005, Distribution free goodness-of-fit tests for linear processes. *Annals of Statistics **33,** 2568-2609.*

[13] Diblasi, A. and Bowman, A.W., 2001, On the use of the variogram in checking the independence in spatial data. *Biometrics* **57**, 211-218.

[14] Durbin, J., 1973, Weak convergence of the sample distribution function when parameters are estimated. *Annals of Statistics* 1, 274-290.

[15] Durbin, J., Knott, M. and Taylor, C.C., 1975, Components of Cramér-von Mises statistics II. *Journal of the Royal Statistical Society, Ser. B* **37**, 216-237.

[16] Eubank, R.L. and LaRicca, V., 1992, Asymptotic comparison of Cramér-von Mises and nonparametric techniques for testing goodness-of-fit. *Annals of Statistics* **20,** 2071-2086.

[17] Franke, J. and Härdle, W., 1992, On bootstrapping kernel spectral estimate. *Annals of Statistics* **20**, 121-145.

[18] Giraitis, L., Hidalgo, J. and Robinson, P.M.,2001, Gaussian estimation of parametric spectral density with unknown pole. *Annals of Statistics* **29**, 987-1023.

[19] Grenander, U., 1950, Stochastic processes and statistical inference. *Ark. Mat.,* **1,** 195-277.

[20] Grenander, U. and Rosenblatt, M., 1957, *Statistical Analysis of Stationary Time Series.* John Wiley, New York.

[21] Guyon, X., 1982a, Parameter estimation for a stationary process on a d-dimension lattice. *Biometrika* **69**, 95-105.

[22] Guyon, X., 1982b, Random fields on a network. Springer-Verlag. Berlin.

[23] Hannan, E.J., 1973, The asymptotic theory of linear time series models. *Journal of Applied Probability* **10**, 130-145.

[24] Hainz, G. and Dahlhaus, R., 2000,, Spectral domain bootstrap tests for stationary time series. Preprint.

[25] Hidalgo, J., 2003, An alternative bootstrap to moving blocks for time series regression models. *Journal of Econometrics* **117**, 369-399.

[26] Hidalgo, J., 2007, Uniform confidence bands for kernel regression estimates with dependent data. Forthcoming in *Journal of Econometrics.*

[27] Hidalgo, J. and Lazarova, S., 2007, Bootstrap in time series regression models. Preprint.

[28] Hidalgo, J. and Kreiss, J.P.,2006, Bootstrap specification tests for linear covariance stationary processes. *Journal of Econometrics* **133**, 807-839.
[29] Hong, Y., 1996, Consistent testing for serial-correlation of unknown form. *Econometrica* **64,** 837-864.
[30] Ljung, G.M. and Box G.E.P., 1978, On a measure of lack of fit in time series models. *Biometrika* **65**, 297-303.
[31] Neyman, J., 1937, Smooth test for goodness of fit. *Skan. Aktuarietidskr.* **20**, 150-199.
[32] Paparoditis, E., 2000, Spectral density based goodness-of-fit tests for time series models. *Scandinavian Journal of Statistics* **27**, 143-176.
[33] Robinson, P.M. and Vidal-Sanz, J., 2006, Modified Whittle estimation of multilateral models on a lattice. *Journal of Multivariate Analysis* **60**, 234-349.
[34] Stute, W., 1997, Nonparametric model checks for regression. *Annals of Statistics* **25**, 613-641.
[35] Whittle, P., 1954, On stationary processes in the plane. *Biometrika* **41**, 434-449.

Economics Department, London School of Economics, Houghton Street, London WC2A 2AE, U.K.

E-mail address: f.j.hidalgo@lse.ac.uk